

1. INTRODUCTION TO THE HIGH SCHOOL TRANSCRIPT STUDY

■ WHAT IS THE HIGH SCHOOL TRANSCRIPT STUDY?

The 1998 High School Transcript Study (HSTS) was conducted by Westat for the U.S. Department of Education's National Center for Education Statistics (NCES). Thousands of transcripts of students who graduated from public and nonpublic high schools were collected from a nationally representative sample of schools. This study provides the Department of Education and other educational policymakers with information regarding current course offerings and course-taking patterns in the Nation's secondary schools. In addition, it provides information on the relationship of student course-taking patterns to achievement as measured by the National Assessment of Educational Progress (NAEP), a federally-funded, ongoing, periodic assessment of educational achievement in the Nation's schools.

Since similar studies were conducted of the course-taking patterns of graduates through the years, changes in these patterns can be studied and compared. Five studies involving the collection of transcripts from high school graduates have been conducted since 1982. These studies consisted of:

Table 1–1. High school transcript studies

Study	Year conducted	Approximate number of transcripts collected
High School and Beyond	1982	12,000
High School Transcript Study	1987	25,000
High School Transcript Study	1990	21,000
High School Transcript Study	1994	25,000
High School Transcript Study	1998	25,000

■ Is participation in HSTS voluntary? Are the data confidential? Are students names or other identifiers available?

The High School Transcript Study is not an exam and requires no preparatory work or time from the students. Students transcripts are collected for the sample of students selected by Westat field workers for the NAEP assessment. All NAEP and HSTS activities are voluntary, so students can refuse to participate in the assessment and/or have a copy of their transcript collected by NAEP staff.

As in NAEP activities, Westat follows the guidance of schools regarding whether or not to inform parents or obtain signed or implied parental consent. Generally, schools do not require parental or

student notification or consent for the HSTS because there is no burden placed on the student. However, if a school requires that students and/or parents be notified or that consent be obtained, Westat complies with that request.

The data obtained from the transcript study are kept strictly confidential. Student names and any other identifiable information are deleted from the copies of the transcripts before these materials leave the schools. Furthermore, in schools that are linked to NAEP, each student receives a NAEP student ID that is also used in the HSTS. The list that links the student's name with that NAEP ID remains in the school. Westat does not have access to that list and cannot re-create it if it is lost.

The data files that Westat supplies (both the restricted and the public use files)¹ do not contain the students' names or other unique identifiers. Data files do contain the students' NAEP ID, which enables researchers to link the transcript data to the NAEP data, but Westat follows NCES' strict procedures regarding the confidentiality of data files.

For more information, please refer to Chapter 4.

■ **What are HSTS's procedures for collecting data?**

The field workers for the 1998 High School Transcript Study were drawn from the pool of NAEP field supervisors and were trained in the data collection procedures.

Eligible schools participating in NAEP were informed about the 1998 HSTS when they received information about NAEP. Schools were provided with information about participation in the HSTS, including procedures that would be used to ensure confidentiality of the data, and the amount and nature of school staff time required for HSTS participation.

Minimal staff time is involved in completing some forms, as well as some clerical effort. Whenever possible, Westat staff assists in these efforts. For eligible schools that agreed to cooperate, students sampled for NAEP were included in the HSTS sample. A brightly-colored Disclosure Notice was placed in their folder both to alert the school personnel that information contained in the student's folder

¹ There are two versions of the 1998 High School Transcript Study data files: the restricted use data files and the public use data files. For a description of those files, please refer to Chapter 7.

would be used for the HSTS and to serve as a visible marker for identifying the folders of selected students to facilitate finding their transcript later.

Initial HSTS information requested from schools and collected by field workers at the time of the NAEP assessment included information which they were asked to provide on the School Information Form (SIF). Other requested information included copies of their school's course catalogs for the four most recent school years, including 1997-98, and three sample transcripts. They were also asked to provide a complete transcript for each graduate in the HSTS sample as soon as graduation information was posted on the transcripts. Information provided on the SIF indicated the appropriate date for the field workers to obtain the transcripts. When completing the SIF, field worker also gathered general information about class periods, credits, graduation requirements, and other aspects of school policy. Sometimes this information was documented in the course catalog and at other times in a separate school policy document.

Field workers filled out checklists for the materials they obtained. These checklists served two purposes:

1. They guided field workers in obtaining materials with the maximum amount of information possible that would be useful in the HSTS.
2. They provided Westat staff with a quick way to review the materials, so that they could request additional information if needed.

This information was collected in visits to the school prior to and at the time of the assessment. When all the information had been collected, it was forwarded to Westat. For schools that did not participate in NAEP but were agreeable to taking part in the HSTS, contact was made near the end of the 1997-98 school year and the same information was collected once the students' final data were posted on their transcripts.

For more information, and a detailed description of the process of obtaining materials for the HSTS, please refer to Chapter 4.

■ **What contextual background data does HSTS provide (a) from questionnaires; (b) from transcripts; and (c) from school-level information?**

The High School Transcript Study provides data that comes from the NAEP questionnaires, the high school transcripts, and the school level forms filled out by a school coordinator or counselor.

QUESTIONNAIRES

- **School Questionnaire:** The School Questionnaire (see Appendix A) is a 54-item questionnaire that collects information about school, teacher, and home factors that might relate to student achievement. It was completed by a school official (usually the principal) as part of NAEP for the NAEP participating schools.
- **Students with Disabilities/Limited English Proficiency (SD/LEP) Questionnaire:** Prior to 1996, the questionnaire that collected information from school staff about students with disabilities and students with limited English proficiency was called the Individualized Education Plan/Limited English Proficiency (IEP/LEP) Questionnaire. It was re-titled as the SD/LEP Questionnaire in 1996. The SD/LEP Questionnaire was completed for students sampled for NAEP and identified by the school as having a disability and/or limited English proficiency. Schools were asked to have the person most knowledgeable about a student complete the questionnaire. In large schools, this person was typically a counselor, a special education teacher, or a teacher of English as a Second Language. In smaller schools, this person was typically a classroom teacher.

For schools participating in the 1998 NAEP, the SD/LEP Questionnaires were collected as part of the NAEP procedures. Questions 1 and 2 were used to determine which section(s) of the questionnaire should be completed. Part A (questions 3 through 19) was answered for a student with a disability. Part B of the questionnaire (questions 20 through 41) was completed for an LEP student. If a student was classified as both SD and LEP, the entire questionnaire was completed. A copy of the questionnaire is included as Appendix C.

TRANSCRIPTS

The student transcripts provide data that is coded and entered into the data system by trained coders. This data include the following:

- Date student enrolled in high school;
- Date student graduated;
- Rank in class;
- Size of class;
- Grade Point Average (GPA);
- Days absent each year;
- Other standardized test scores and honors (where available);

- List of courses taken in high school, including the grades received and the number of credits received for each course; and
- Total number of credits received and, in many cases, total number of credits attempted.

SCHOOL-LEVEL INFORMATION

- **Transcript Request Form (TRF):** When graduation information was posted on the transcripts, field workers returned to the schools to obtain the requested transcripts. For each NAEP school, the field worker was given a Transcript Request Form (See Chapter 4). In addition to the NAEP ID, it contained columns for entering graduation status (Exit Status) and the student's gender, birth month and year, race/ethnicity, SD status, LEP status, Title 1 services receipt, and National School Lunch Program participation. Data available from NAEP files (NAEP ID and demographic variables) were preprinted on the form. After completing the form, any personal identifiers were removed from the Transcript Request Forms. Westat did not include the students' names on the TRFs.
- **School Information Form (SIF):** The SIF (see Appendix B) was completed by the field worker or a school staff member or sometimes by both. The completed SIF contained information about the school in general, about sources of information within the school (if needed to complete HSTS data collection), course description materials, significant changes in course offerings in the past four years, graduation requirements and grading practices at the school, and about the format of the school's transcripts. The field workers were instructed to fill out the SIF completely, or to indicate clearly on the SIF where the requested information could be found in the other materials provided by the school.
- **School-Level Catalog or Course Lists:** If a school provided catalogs of course offerings for the four years that the seniors attended the school (as requested), data entry personnel entered a list of all course titles appearing in the catalogs. A concerted effort was made to standardize the format of titles. About 75 percent of the schools provided more than one year's catalog. Catalogs from all years received were used to determine whether there were significant changes over the four years. A curriculum specialist selected the portions of each catalog to be used so that they excluded sections on programs that students could take only by attending another school in the district, courses taken at night, and so on. The specialist included programs from previous years that were not listed in the current catalog but were offered during the period when students in the HSTS attended the school. These titles were entered in the order of their appearance in the catalogs.

For more information, please refer to Chapters 4 and 5.

- **What is the Transcript Request Form (TRF) and why is it so important? How is the TRF obtained and what information does it contain?**

When graduation information was posted on transcripts (the date that it would be available was provided by the school on the School Information Form), a field worker returned to the school to obtain the requested transcripts. For each NAEP school, the field worker was given a Transcript Request Form (TRF), Version 1 (see Exhibit 4-7). In addition to Student Name and NAEP ID, it contained columns for the students' graduation status, gender, birth month and year, race/ethnicity, SD status, LEP status, Title 1 services receipt, and National School Lunch Program participation. Data available from NAEP files (NAEP ID and demographic variables) were preprinted on the form. The completed TRFs contained the following information:

- **Student Name** – Since names were never removed from the school, this column was blank when the TRF's were printed. The field worker first recorded the first name, middle initial, and last name of each assessed, absent, or excluded student listed on the NAEP Administration Schedule. The names were recorded only to ensure that the correct student folders were used.
- **NAEP ID** – The 10-digit NAEP assessment booklet numbers, or SD/LEP questionnaire numbers for students excluded from the 1998 assessment were preprinted in ID order. This column on the TRF identified all students for whom transcripts were needed.
- **Exit Status** – Using information provided by the school, field workers assigned one of the following codes to describe each student's outcome at the school. The Exit Status codes are defined on page 1-16.
- **Birthdate, Gender, and Race/Ethnicity** – Demographic information was generally preprinted for each sampled student. If not preprinted, it was recorded from the NAEP Administration Schedule. If the school informed a field worker that some of this information was incorrect, the field worker entered the correct information on the TRF.
- **SD and LEP Status** – For each student, it was recorded whether or not the student was classified by the school as SD and/or LEP.
- **National School Lunch Program (NSLP) and Title 1** – Yes or No for participation in each of these programs.
- **Transcript Received** – Field workers checked this column to document that the transcript for a given student had been received.

Once the Transcript Request Form was completed, the field worker filled out a summary box at the top of the form and requested transcripts according to the procedures set forth by the school. The

Disclosure Notice placed in students' folders at the time of the NAEP assessment helped to facilitate transcript collection in participating NAEP schools.

Once the field worker filled in the names of the students, some schools were able to access an electronic data file and copy the transcripts. In other schools, the transcripts were manually pulled from their folders and photocopied at the school.

Once the transcripts were provided, the field worker reviewed them to ensure that a transcript was received for each 12th grade student who was selected for the NAEP assessment, whether or not that student had graduated. (Non-graduates were removed from the files at a later stage.) The field worker then checked each transcript for eligibility, understandability (e.g., are all the codes on it defined on the transcript or explained in the SIF?), and completeness. The field worker labeled each transcript with preprinted labels containing the School ID and the NAEP ID for the student and completed a Documentation of Missing Transcripts form to explain any omissions.

After the field worker collected and recorded all the information required on the sampled students and reviewed the transcripts for completeness and accuracy, he or she prepared the transcripts for transmittal to Westat. This procedure involved "masking" all personally identifiable information where it appeared on each transcript, using a broad felt tip marker or correction tape to line through or cover all identifiers.

Personal identifiers were also removed from the Transcript Request Forms. Before sending the TRFs back to Westat, the field worker cut off the portion that included the students' names, to comply with confidentiality provisions. The portion with the names was left in the school's NAEP folder.

For further information, please refer to Chapter 4.

■ **What is a Course Catalog and how is it obtained? What is the difference between the different types of catalogs, and which one is used in the data processing step?**

A Course Catalog is a list of all the courses the school offers and their descriptions. Such catalogs are generally published each year, and are used for accurately matching the CSSC codes used in the HSTS with the course title, obtained from the transcript.

Field workers request these Course Catalogs when they first contact a school, then collect them when they visit the school for sampling. The Course Catalogs are carefully reviewed at the school. Field workers verify that the catalog contains all the courses that the 12th graders of that year have taken in that school, including vocational, remedial, honors, special education, off-campus courses, or courses taught in a language other than English. If any course listings were not in the catalog, every effort is made to obtain additional information from school personnel to document the existence of such courses and to describe them. After that review, the Course Catalogs are sent to Westat.

In most cases, the current Course Catalog and the ones from the three preceding years are collected. This allows Westat to track any changes in course offerings or in the curriculum in the four years the graduates attended high school. It also allows the catalog coders to review any course title on the transcript and accurately match it to a description in the catalog, even if the curriculum or the course titles have changed during those four years.

Based on Westat's experience in coding course catalogs from this and previous High School Transcript Studies, five types of course catalogs have been identified:

1. A school-level catalog providing course titles and descriptions;
2. A district-level catalog, which indicates which courses were offered at the HSTS participating school;
3. A course list by department that includes general descriptions of course offerings by department;
4. A school-level course list without descriptions; or
5. A district-level catalog without any indication of which courses were offered in specific schools.

Westat uses the highest-level catalog available.

For further details, please refer to Chapter 4.

■ **Who codes the catalogs? What special requirements are needed from the coders? How are they trained?**

In order to code the school catalogs, Westat hired a staff of skilled personnel who had an extensive background in education, mainly teachers and counselors, and who were familiar with schools'

curriculum and the education system. These staff members were trained to familiarize themselves with the CSSC coding scheme and the variety of ways that a course could be coded. For several days, they were given exercises and tasks to ensure that they could code a course title with the appropriate CSSC code.

To ensure consistency and quality, catalog coding decisions were based on a basic set of coding principles and procedures. First, the catalog coder reviewed a school catalog “holistically” to ascertain ways that course levels, special education, and other special programs were designated. He or she looked for sequences of courses, descriptions of programs, requirements, credits awarded, or other information provided to obtain a general view of the curriculum. Then, using the CACE system (Computer Aided Coding and Editing), the coder looked at each course title, found it in the catalog, and read whatever description was available. The coder then selected the best CSSC code for the course. Wherever possible, the catalog coder selected codes based on a course description rather than on title.

After selecting the CSSC code, the coder reviewed the flags for that course and edited them as needed. If the coder found courses in the CACE catalog listing that should not be there, they could be deleted. Similarly, if the coder found that a course was missing from the CACE listing of catalog titles, it was added to the list and coded. After the coder finished coding the regular education courses for a school, the special education expert coded all special education courses.

For the specific steps of the coding procedure please refer to Chapter 5.

■ **How are the data entered from the transcripts?**

Westat processed the data from the 1998 High School Transcript Study (HSTS) along three simultaneous paths as follows:

1. The Student Sampling Information System;
2. The Computer Assisted Data Entry System; and
3. The Computer Assisted Coding and Editing System.

With the exception of the transcripts and the course catalogs, some data entered in each system were collected by Westat field personnel and some data had already been assembled for NAEP into data files by the Educational Testing Service (ETS). Westat staff obtained the relevant NAEP data

files from ETS and merged them with the HSTS data collected from non-NAEP-participating schools. As described below, appropriate checks were made to ensure that only one set of data was entered for a school or a student, and procedures were developed to resolve inconsistencies among the data sources.

When entering and cleaning the data for the study, we performed the following tasks:

- Establishing student ID control lists;
- Entering transcript data;
- Coding course catalogs;
- Matching transcript titles to catalog titles;
- Standardizing credits and grades; and
- Performing quality control checks.

These steps involved the entry and coding of the students' transcripts and the schools' course catalogs, as well as matching the courses on the coded catalogs to the courses on the transcripts.

Each of these steps is described in detail in Sections 5-1 through 5-6 of Chapter 5.

HOW IS THE HIGH SCHOOL TRANSCRIPT STUDY RELATED TO THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS?

■ How is HSTS related to NAEP?

The High School Transcript Study is conducted in conjunction with the National Assessment of Educational Progress (NAEP); both are federally-funded and completed by Westat for the U.S. Department of Education's National Center for Education Statistics. The 1998 HSTS was designed to allow an analysis of the course-taking patterns of students who graduated from American public and nonpublic high schools in 1998. It was further designed so that data on students' course-taking patterns can be linked to the 1998 NAEP assessment results. Since studies similar to the 1998 HSTS were

conducted on 1982, 1987, 1990, and 1994 graduates, changes in these patterns and relationships to NAEP performance in these years can also be studied.²

NAEP provides the HSTS with data of assessments in different subjects. For the 1998 HSTS, NAEP proficiency estimates of reading, writing, and civics were provided. In 1994, history, geography and reading assessments were conducted, and proficiency estimates in those subjects were provided.

For a comprehensive description of the HSTS and NAEP please refer to Chapter 2.

■ **How are the samples of schools and students in NAEP related to the HSTS samples?**

In order to maintain as many links as possible with NAEP scores, where schools refusing to participate in NAEP were replaced by substitute schools, the substitute schools, not the refusals, were asked to participate in the HSTS. Of the 322 schools in the original sample, 264 original/substitute schools participated in the HSTS survey, of which 241 were originally sampled. Of the 264 participating schools, 232 schools cooperated with both HSTS and NAEP and the links for the students were maintained, 10 schools cooperated with HSTS and NAEP but the links for the students were not maintained, and 22 schools cooperated with HSTS but not with NAEP.

A total of 28,764 students were selected for inclusion in the HSTS. Of these, 27,183 students were from schools that maintained their NAEP administration schedules and were identified by their NAEP booklet numbers. Another 500 students were from schools that participated in NAEP but had lost the link between student names and NAEP booklet numbers, and 1,081 were from schools that did not participate in NAEP.

Because sampling was performed in most schools prior to graduation, not all sampled students were, in fact, graduates. Only graduates, however, were eligible for the transcript study. We were able to determine that of the 28,764 students in the sample, 25,248 actually graduated and that 3,328 did not. Of the remaining 188 students, we imputed 174 as graduates and 14 as not. Thus, from the 25,422 graduates 24,218 transcripts were collected and processed.

² The 1987, 1990 and 1994 transcript data were collected by Westat in coordination with the 1987, 1990, and 1994 NAEP (Thorne et al., 1989; Legum et al., 1993; Legum et al., 1997). The 1982 data were collected by the National Opinion Research Center as part of the High School and Beyond project (Jones et al., 1983a).

For further information regarding this topic, please refer to Chapter 3.

HOW CAN THE DATA FROM THE HIGH SCHOOL TRANSCRIPT STUDY BE USED?

■ Can the 1998 HSTS results be compared to other transcript studies?

Since studies were conducted of course-taking patterns of 1982, 1987, 1990, 1994 and 1998 graduates, one research objective was to study changes in these patterns. Another research objective was to compare course-taking patterns to study results on the 1998 National Assessment of Educational Progress.

The 1998 High School Transcript Study used a complex sample design with the goal of securing a sample from which estimates of population and sub-population characteristics could be obtained with reasonably high precision (in other words, low sampling variability). At the same time, it was necessary that the sample be economically and operationally feasible to obtain. The resulting design requires that the user of the HSTS data utilize sampling weights to ensure valid analysis of the transcript data.

The samples for the five studies are roughly comparable. The weighting techniques used across studies are also comparable. However, there are some differences that warrant notice when using these tables.

The 1982 sample was drawn as part of the first follow-up to the High School and Beyond longitudinal study. The 1987, 1990, 1994, and 1998 samples were drawn as part of the corresponding NAEP samples. One result of this difference is that the 1982 study, because of its longitudinal nature, had more opportunity to obtain demographic information. On the other hand, because students repeated years, transferred to different schools, or dropped out of school before their senior year, there was also a greater probability that final transcripts showing four years of high school could not be obtained for these students.

The samples were drawn at different points in the students' high school careers. The 1982 students were sampled when they were sophomores and were followed when they transferred to new schools. The 1987 students were sampled when they were juniors, but no attempt was made to follow

them if they left school. The 1990, 1994, and 1998 students were sampled in their senior year. Thus the 1987 study sample, unlike the others, has no students who transferred into their school during their senior year.

All five samples used a multi-stage, stratified, and clustered design. There are differential rates of oversampling among the studies to reflect special interests. For instance, the 1987 study oversampled students with disabilities and the 1994 and 1998 NAEP assessments oversampled minority students.

Westat performed all the variance estimations using the jackknife procedure. Because the number of replicates used in the 1990 study was greater than in the earlier studies, the variance estimates for 1990 are somewhat more precise than in the earlier studies. Similarly, because the number of replicates used in the 1994 and 1998 studies were greater than in the earlier studies, the variance estimates for 1994 and 1998 are even more precise. Note that the 1982 sample consisted of considerably fewer transcripts than in later years. The number of schools involved, however, was considerably greater. The estimates tend to have comparable sampling errors across years, despite the differences in the number of transcripts sampled. The sampling errors, in fact, are often smaller for 1982 estimates. In other words, the design effects for years other than 1982 were considerably greater than for 1982, more than offsetting the effects of the larger sample size of transcripts for those other years.

The sample sizes differ in the five studies and are summarized in the table below.

Table 1–2. Sample sizes for the high school transcript studies

Sample size	1998	1994	1990	1987	1982
Schools in the original sample	322	379	379	497	1,882
Schools represented in the tables	264	340	330	429	947
Students in the original sample	28,764	28,715	23,270	35,180	18,427
Transcripts represented in the tables	24,218	24,120	21,158	25,054	12,275
Average number of transcripts per school	96.3	73.8	64.1	58.2	12.9

The data tables for the 1998 HSTS are presented in Appendix A of the Tabulations Report. Appendix B of the Tabulations Report contains a listing of the categories (stubs) used as row labels in the tables and the CSSC codes associated with each category.

For more information about the 1998 tabulations, and the comparison between the different studies, please refer to Chapter 1 of the Tabulations Report. Please refer to Chapter 2 of this guide for a comprehensive description of the NAEP study.

■ **What is a weight and how is it determined?**

The High School Transcript Study sampled almost 29,000 students from 264 schools. To make valid inferences about the entire population of graduated grade 12 students from the sample of student transcripts that was collected, it is necessary to use the sampling weights. The weights reflect the probability sampling scheme used to arrive at the sample of students for whom transcripts were requested. The HSTS weights were constructed without regard to the NAEP participation or nonparticipation status of schools and students. The weights also reflect the impact of sample nonresponse at the school and the student levels, making adjustments for these groups to decrease the potential bias that might arise through differential nonresponse across population subgroups. Finally, improvements to the precision of weighted estimates result from the application of poststratification factors to the sample weights.

Student transcript data were weighted for the purpose of making estimates of course-taking by high school graduates nationwide. The final weight attached to an individual student record reflected two major aspects of the sample design and the population being surveyed. The first component, the base weight, was used to expand sample results to represent the total population and reflected the probability of selection in the sample. The second component, the adjustment of the base weight to account for nonresponse within the sample, is implemented to ensure that the resulting survey estimates of certain characteristics (race/ethnicity, size of community, and region) conformed to those estimates known reliably from external sources.

Weights, developed using the procedures described in Chapter 3 of the Tabulations, are contained in the Student File and the Linked Weights File. Westat has provided the final student weight (FINSTUWT) in the Student File and the final usable linked weight (FINLNKWT) in the Linked Weights File so that data analyses can be weighted up to national totals. The final student weight should be used in analyses involving only transcript data. The weights in the Linked Weights File should be used in analyses involving both transcript data and data obtained from NAEP data files.

For further information, please refer to Chapter 3 of the Tabulations Report.

- **Why are there two general sets of weights (linked and non-linked weights) for HSTS?**

The linked weights must be used whenever the analysis involves NAEP data. There are files containing linked weights for reading, writing, and civics, the subjects in which students were assessed in 1998 NAEP. The linked weights were created to analyze each NAEP subject separately. Conversely, the non-linked weights must be used when analyzing transcript data only (i.e., without regard to NAEP data). The student file, which lists all HSTS students, contains these non-linked weights. One difference between the processes for creating linked and non-linked weights is in the treatment of nonresponse. The linked weights are adjusted to account for nonrespondents, where nonrespondents are eligible students with incomplete transcripts or eligible students that were absent in NAEP. The non-linked weights are adjusted to account for nonrespondents, where nonrespondents are eligible students with incomplete transcripts.

For more information about the linked weights, please refer to Chapter 3.

- **What is the PSU?**

The HSTS used a subsample of primary sampling units (PSUs) and schools from the 1998 NAEP assessment for grade 12 students. The HSTS used the NAEP target sample of students in these subsampled schools. Chapter 3 describes aspects of the selection of PSUs, schools, and students that are specific to the transcript study. The purposes of the 1998 High School Transcript Study (HSTS) were to gather data on a nationally representative sample of students who graduated from American public and nonpublic high schools in 1998 that could be linked to NAEP results. For the HSTS sample of students to be as representative as possible, it included a subsample of NAEP PSUs, and, subsequently, subsampled schools with 12th grades within the PSUs that were selected for NAEP, regardless of whether they participated in NAEP. A representative sample of students was included from each school. When possible, the students selected for the transcript study were the same as those selected for NAEP. When this selection was not possible, a systematic sample of students was drawn from the school.

For further information, please refer to Chapters 2 and 3.

■ **How is a student given a unique HSTS ID?**

The 1998 High School Transcript Study involved collecting, processing, and analyzing nearly 29,000 transcripts from 264 high schools nationwide. In order to accurately process each of these students' transcripts, a unique student ID was necessary.

The HSTS school sample was a sub-sample of NAEP. Each one of the schools participating in the HSTS had a unique 3-digit Primary Sampling Unit (PSU) ID and a 3-digit School ID.³ Each school then had a unique 6-digit ID.

In NAEP-linked schools, i.e. schools that retained their link to the NAEP assessment, each student received a unique 10-digit booklet ID. This ID was unique throughout the HSTS. A different procedure was required for schools that had not retained their materials which linked selected students to their specific IDs or had not participated in NAEP. In those schools, a new sample of students was selected and students were assigned IDs that ranged from 990–0000001 to 990–0000060. This scheme meant that the student IDs were unique within a school, but not within the entire study.

In order to achieve unique student IDs within the entire study, the school's 6-digit unique ID was concatenated to the 10-digit student ID. This assured that each student received a unique ID across the entire study.

■ **What is an Exit Status and how is it used? Why are there more Exit Statuses in 1998 than in previous years?**

The Exit Status is a code that describes the type of diploma the student received. Using information provided by the school, field workers assigned one of the following codes to describe each student's outcome at the school:

- a. Graduated with a standard diploma;
- b. Graduated with an honors diploma;
- c. Received a diploma with special education adjustments;
- d. Received a certificate of attendance;

³ The School ID is a 3-digit ID to which a fourth control digit is added. In many of the reports, Westat included this fourth digit, but for the purpose of obtaining a unique student ID, this digit was dropped.

- e. Still enrolled in this school;
- f. Dropped out;
- g. Other, such as transferred, GED, or unknown;
- h. Out of Scope; or
- i. Completed course requirements but did not pass required tests.

In some cases, the Exit Status was determined directly from the transcripts and sometimes it was provided by other sources at the school. The Exit Status was recorded on the Transcript Request Form (TRF) and later used to verify that the student indeed graduated and that his/her transcript was eligible for the study. It also provided information about whether or not to include the transcript in the tabulation process. In a few cases, Westat discovered that a student had not actually graduated and changed the Exit Status accordingly.

In 1998, two new Exit Statuses that did not exist previously (H and I) were added to the list. Exit Status H was added to address cases in which the student was excluded from the study, such as students who graduated during the study year (1998) but who had been attending high school for more than five years. Exit Status I was added to describe a case where a student fulfilled the school's requirements for graduation, but did not pass a state exam that made him/her eligible for a graduation diploma.

For more information about the Exit Status, please refer to Chapters 4 and 5.

■ **What is the CSSC Code and how is it used? Are there any other coding systems that are being used in similar studies?**

To compare transcripts from different schools, it is necessary to code each of the courses entered from the transcripts using a common course coding system. The coding system employed for this purpose was a modification of the Classification of Secondary School Courses (CSSC). The CSSC, which contains approximately 2,200 course codes, is a modification of the Classification of Instructional Program (CIP) that is used for classifying college courses. Both systems (CIP and CSSC) use a three-level, 6-digit system for classifying courses. The CSSC uses the same first two levels as the CIP, which

are represented by the first four digits of each code.⁴ The third level of the CSSC (the fifth and sixth digits of the course code) is unique to the CSSC and represents specific high school courses.

With over 2,200 codes in the CSSC, it is neither practical nor desirable to include estimates of each possible code in each of the tables. Although estimates are provided for each of the codes that appear in the transcripts in the final table in Appendix A of the Tabulations, it is often more useful to analyze the courses in larger groups such as English, Social Studies, Math, or Science. There is also interest in finer divisions of these groups such as Biology, Chemistry, and Physics. The subject area taxonomy that is presented in Appendix B of the Tabulations provides the structure for grouping the courses.

The 1987 High School Transcript Study developed a taxonomy used for the 1987 High School Transcript Study Tabulations. This taxonomy, which is documented in the 1987 Tabulations, was developed with an emphasis on strictly limiting the content of “academic” categories to academic courses. It was applied to data from the 1982 High School and Beyond (HS&B) First Follow-up Study and the 1987 HSTS data. Both of these data sets were coded using the CSSC. The 1990 High School Transcript Study used a slightly expanded version of the same taxonomy in its reports.⁵

The Secondary School Taxonomy (SST) was originally developed in the late 1980s.⁶ In addition to the HS&B and 1987 HSTS files, variants of the SST were applied to files produced by the Educational Testing Service Study of Academic Prediction of Growth (1969) and the National Longitudinal Study-Youth Cohort (1975-1982), which were coded using unique classification schemes which were not fully compatible with the CSSC. The SST was developed under the auspices of the National Assessment of Vocational Education (NAVE) and was subject to extensive review by vocational and academic educators and researchers, NAVE staff, and contractor staff. Although there is broad agreement between the taxonomy developed for the 1987 HSTS and the SST, the SST has a less purely academic emphasis and a more richly defined group of vocational education categories.

⁴ Actually, the CSSC uses the first two levels of the CIP as it existed in 1982. The CIP has undergone some modification since then. In addition, three sets of codes at the top level have been added to the CSSC to provide a means of classifying courses specifically designed for students with disabilities.

⁵ The 1990 study added 18 new codes to the CSSC and to the taxonomy. The full taxonomy is documented in both *The 1990 High School Transcript Study Tabulations* and *The 1990 High School Transcript Study Data File User's Manual*.

⁶ A description of the development of the SST is provided in Gifford, Hoachlander, and Tuma (1994), *The Secondary School Taxonomy Final Report*.

Since most recent NCES publications which have analyzed transcript data have used the SST, it was adapted for use in the 1994 tabulations, and this adaptation has been carried over to the 1998 report. The SST is, however, limited in that it contains only the CSSC codes found in the data sets which it was designed to analyze. For this reason, the SST was expanded in 1994 to include all currently defined CSSC codes.⁷ The version of the Secondary School Taxonomy used in these tables also differs from the version used in studies before 1994 in two other respects:

- Some additional categories have been added. These have not changed the definition of any of the existing categories.
- Drama and Dance have been separated into two categories. This split is consistent with the reporting level in the previous High School Transcript Studies. Since these two values are always reported adjacent to each other, they can easily be added together to determine the corresponding combined category.

Because the SST assigns courses differently to academic and vocational categories, analyses based on the SST report larger numbers of students following vocational curricula and fewer following academic curricula than the taxonomy used in the 1987 and 1990 transcript studies. For example, the 1990 HSTS classified 75 percent of 1990 graduates⁸ in academic programs and 6 percent in vocational programs (1990 HSTS Tabulations, Table 1), while the current study classifies 69.6 percent of 1990 graduates in academic programs and 8.0 percent in vocational programs (Table 2).

One other feature of the SST that should be kept in mind when interpreting these tables is that it classifies English as a Second Language (ESL) courses as Foreign Language rather than English courses. This classification has the effect of lowering the number of students who appear to satisfy the recommendation of completing four years of English. It also has the effect of increasing the apparent number of Foreign Language courses completed and lowering the correlations of number of years of Foreign Language completed with each set of the NAEP proficiency scores. These effects are particularly noticeable among Hispanic graduates.

⁷ In addition to the studies cited earlier in this section, the National Education Longitudinal Study (NELS) of 1988 Second Follow-Up: Transcript Component collected transcripts from high school graduates and coded them using the CSSC. The students in the transcript component of the NELS study graduated from high school in 1992. Researchers at National Opinion Research Center, which conducted the study for NCES, have informed us that they were able to use the CSSC codes in the 1990 version of the CSSC and did not need to add any additional codes.

⁸ Legum, Stanley; Caldwell, Nancy; Goksel, Huseyin; Haynes, Jacqueline; Hynson, Charles; Rust, Keith; Blecher, Nina. *The 1990 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 1990, 1987, and 1982 High School Graduates*. U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, NCES 93-423, Washington, DC, April 1993.

■ **How are codes added to the CSSC? Are they ever deleted?**

The high school curriculum may change each year or every few years. New courses are added, old courses are taken out of the curriculum, and some courses are combined with others to produce new courses. The CSSC code list contains over 2200 codes and descriptions of courses offered by high schools nationwide. For every HSTS, the need arises to examine the list of CSSC codes and decide whether all the courses that were offered in that particular year have a matching CSSC code that can adequately describe it. In 1994, 12 new CSSC codes were added to the list. In 1998, the computer science curriculum changed dramatically. New courses such as Web Design, Java Programming, and C++ Programming were added, courses that did not exist previously. Also, many courses that were labeled as honor courses in the past were reclassified as AP courses. Many IB (International Baccalaureate) courses were added as well. In all, a total of 83 new or revised codes were added to the CSSC in 1998.

Highly trained coders are hired to code the school catalogs Westat had received from the field workers. These coders browse through the catalogs and match the appropriate CSSC codes to the courses offered, according to the content and description of the course. If a course that is offered does not have a matching CSSC code in the existing list, the coders write that course description in a special suggestion list. After the catalogs have been reviewed, and all but these courses on the suggestion list have been coded, a Coding Specialist reviews the suggestion list and tries to match these courses to existing CSSC codes. If a course does not have a matching CSSC code, a new CSSC code is generated.

Schools also make changes to their curriculum by dropping courses they had offered in the past. These courses are either dropped completely from the offering list, split into two courses, or are renamed and their course description changed. During the data cleanup stage that Westat performed, duplicate or unneeded CSSC codes were deleted. An example would be a Calculus AP honors course that was redefined and split into Calculus AP and Calculus honors. Both new CSSC codes better describe the offered courses, so that the original CSSC code was no longer needed. Westat made sure that for each CSSC code that was deleted, documentation was supplied and analyses across the HSTS studies could be carried out.

For further information about the CSSC codes, please refer to Chapter 7.

■ **Were there any restrictions on what data appeared in the HSTS tabulations?**

For the 1998 HSTS, Westat attempted to collect high school transcripts from about 29,000 students who graduated from high school in 1998. So that the tables represent students with complete transcripts, students whose transcripts did not include course-by-course data for at least three full years of high school were excluded. To be consistent with other published analyses, Westat adopted the following rules for including and excluding students in the analyses that produced the tables:

1. Both public and nonpublic school students were included.
2. Students with special education diplomas, certificates of attendance, and certificates of completion were excluded. Students who received certificates of completion completed the necessary school requirements for graduation, but failed to successfully complete a required state graduation exam.
3. Students with disabilities (HCFLAG=2 in the HSTS studies) who received regular or honors diplomas (i.e., those who were not screened out by rule 2) were included.
4. Students with zero English credits were excluded.
5. Students with fewer than 16 Carnegie units were excluded.

Some previous studies have excluded students with more than 32 Carnegie units on the grounds that they must have shorter class periods than normal schools and use of their data would inflate our estimates. In the current study, students with more than 32 Carnegie units were not excluded.

In a few cases, Westat discovered that a student had not actually graduated and changed the exit status accordingly. It was also found that some students had earned substantially more credits than were required to graduate. Often these were students who had spent substantial amounts of time in both foreign and American high schools. While they were awarded credit for the foreign courses, they were still required to take an essentially American curriculum in order to obtain the American diploma.

In still other cases it was found that, although a student had fewer credits than were required to graduate, the transcript had all the other attributes of a graduated senior such as four full years of courses, all required courses, a graduation date, grade point average, and class standing. In these cases, if a careful review of the transcript and the data files showed no data entry or coding errors, the transcript was kept in the database with the apparent inconsistency recorded on the transcript, but was not included in the tabulations.

These restrictions reduced the number of 1998 graduates represented in the tables to 24,218. These students attended 264 schools that had previously been sampled for the National Assessment of Educational Progress.

For further information on this topic, please refer to Chapter 2 of the Tabulations Report.

■ **How does one obtain copies of HSTS data files?**

There are two versions of the 1998 High School Transcript Study data files: the restricted use data files and the public use data files. All values in this report are based on the restricted use data files. To ensure the confidentiality of students, data in the School File, Course Offering File, and Transcript File that would identify the state in which a school is located have either been set to missing (as in the FIPS State Code in the School File) or set to generic values (e.g., a course title of “Mississippi History” was set to “State History” in the Course Offerings File). In addition, the number of teachers and enrollment values in the School File and some race/ethnicity values in the Student File have been set to missing. The data in the remaining files are identical in both the restricted use and public use versions.

Because of confidentiality legislation, secondary users who wish to obtain a copy of the restricted-use data files must apply for an NCES restricted data license. If your organization does not already have a restricted data license, you need to obtain a copy of the “NCES Field Restricted Use Data Procedures Manual.” There is a four-page checklist in this document that details the steps involved in obtaining a license. You may request a copy from the following contact person or you may view and download the manual from the NCES web site at <http://nces.ed.gov/statprog/rudman> .”

Cynthia Barton (202) 502-7307
cynthia_barton@ed.gov

If your organization already has a restricted data license, you may only need to have it amended to add any additional datasets or to add additional names as authorized users of the data. Note that, in a college or university setting, only faculty can serve as the primary project officer. Graduate students may be listed as authorized users only.

To obtain a restricted data license (or to amend an existing license), a secondary user generally must send a letter addressed to the Data Security Office, formally requesting the data. The mailing address of the Data Security Office is:

Data Security Office
Department of Education/NCES
1990 K Street NW
Room 9061
Washington, DC 20006

Please include the following information in your request:

- The name of the dataset(s) you wish to use;
- The purpose for the loan of the data;
- The length of time you will need the data; and
- An affidavit of nondisclosure for each person who will have access to the data, promising to keep the data completely confidential.

For other publications of previous High School Transcript Studies and NAEP, please contact Cynthia Barton at the number given above.

Brief summaries of the transcript coding system, the demographic information that was collected on the students in the study, the student data weights, and the data files are provided below. There is more detail on each of these subjects in the following chapters of this report.

The Coding System

To compare transcripts from different schools, it is necessary to code each of the courses entered on the transcripts using a common course coding system. The coding system employed for this purpose was a modification of the Classification of Secondary School Courses (CSSC) (Ludwig et al., 1982). The CSSC, which contains approximately 2,200 course codes, is a modification of the Classification of Instructional Programs (CIP) that is used for classifying college courses (Morgan et al., 1991). Both systems use a three-level, six-digit system for classifying courses. The CSSC uses the same

first two levels as the CIP, which are represented by the first four digits of each code.⁹ The third level of the CSSC (the fifth and sixth digits of the course code) is unique to the CSSC and represents specific high school courses.

The CSSC also uses an additional 1-digit “disability” flag and a 1-digit “sequence” flag. The first flag indicates whether a course is open to all students or is restricted to disabled students. The sequence flag indicates whether a course is part of a sequence of courses and, if so, its place in that sequence. The disability flag was added to the CSSC during the 1987 transcript study. The sequence flag was added during the 1990 study.

During the 1987, 1990, 1994, and 1998 studies, courses appearing on student transcripts were coded to indicate whether they were transfer courses, offered off campus, honors or above grade-level courses, remedial or below grade-level courses, or designed for students with limited English proficiency (LEP) and/or taught in a language other than English. In 1998, courses offered as Advanced Placement or International Baccalaureate courses were coded separately from other honors-level courses, using both new CSSC codes and new flag values. A total of 83 new or revised CSSC codes were added in 1998. In addition to codes for Advanced Placement and International Baccalaureate courses, most new codes reflect changes in course offerings in the technology area.

Course catalogs and related materials and information from the participating schools were used to determine the codes assigned to each course. Grades and credits were also received and entered for each course and standardized into a consistent system.

Student Information

Information gathered for all students included gender, grade level, birth year, birth month, graduation status, race/ethnicity, whether or not the student had a disability (SD) or was limited English proficient, received Title 1 services, or participated in the National School Lunch Program. Also obtained were the date of entry to the school, the graduation date, type of diploma, number of days absent in each of four years (9th grade, 10th grade, 11th grade, and 12th grade), grade point average, and class rank. In

⁹ Actually, the CSSC uses the first two levels of the CIP as it existed in 1982. The CIP has undergone some modification since then. In addition, three sets of codes at the top level have been added to the CSSC to provide a means of classifying courses specifically designed for students with disabilities.

addition, all awards and scores on certain standardized tests taken by each student as reflected on the transcript were listed.

In some cases, more than the basic information was obtained. The following additional information, as reported by school personnel, was collected for students with a disability: grade-level equivalent performance in English and mathematics, proportion of time the student was placed in mainstream and special education classes, type and severity of disability, and type of accommodation(s) provided to the student.

The following additional information, as reported by school personnel, was collected for students with limited English proficiency: English and mathematics grade levels, percentage of the day spent in special language programs, native language, type of specialized instruction, the type of accommodation(s) provided to the student in testing, and the student's ability to speak, understand, read, and write English.

Student Data Weights

Student transcript data were weighted for the purpose of making estimates of course-taking patterns by students in the class of 1998 nationwide. Several sets of weights were created:

- Weights for all eligible sampled students with completed, missing, or unusable transcripts in the transcript study, where “eligible” means that the student graduated in 1998, and “unusable” transcripts are those with less than 75 percent of the credits required by the school to graduate. Weights are set to zero for missing and unusable transcripts.
- Four sets of “linked” weights for NAEP-assessed and excluded students who graduated and for which a usable transcript was obtained. Since students in NAEP were assigned an assessment of a particular subject, separate weights were developed for the students in each subject-specific assessment:
 - NAEP 25-minute writing assessment;
 - NAEP 50-minute writing assessment;
 - NAEP reading assessment;
 - NAEP civics assessment; and

- The NAEP study Assessment of Civics Trend was not analyzed and was not included in the 1998 HSTS.

In each set of weights, the final weight attached to an individual student record reflected two major aspects of the sample design and the population surveyed. The first component, the student base weight, is the reciprocal of the probability of selection into the sample, which takes into account the product of the probability of selecting the primary sampling unit (geographic area), the probability of selecting the school within the primary sampling unit, and the probability of selecting the student within the school. The second component resulted from the adjustment of the student base weight to account for nonresponse within the sample and to ensure that the resulting survey estimates of certain characteristics (race/ethnicity and region) conformed to those known reliably from external sources.

In order to make valid inferences about the entire population of graduated 12th grade students from the sample of student transcripts collected, it is necessary to use the sampling weights. The weights reflect the probability sampling scheme used to arrive at the sample of students for whom transcripts were requested. The HSTS weights were constructed without regard to the NAEP participation/nonparticipation status of schools and students. The weights also reflect the impact of sample nonresponse at the school and the student level, and make weight adjustments to decrease the potential bias that might arise through differential nonresponse across population subgroups. Finally, improvements to the precision of weighted estimates result from the application of poststratification factors to the sample weights.

Data Files

The study has produced a set of eight data files that are available on public use data sets (with some additional information available on a restricted use basis):

- The Master CSSC File – The Classification of Secondary School Courses (CSSC), including all modifications made to the original (1982) CSSC during the 1987, 1990, 1994 and 1998 transcript studies. This file has separate variables for the CSSC code, the disability flag, the sequence flag, and the course title.
- The Course Offerings File – Provides a comprehensive listing of the courses offered in the schools included in the study. A CSSC code is associated with each course title.
- The School File – Provides detailed information on the schools from which the students were sampled.

- The Student File – Provides demographic information on all students in the study, as well as sampling weights and summaries of their course-taking histories.
- The Four Subject-Level Linked Weights Files – Provides weights for use when performing analyses relating transcript data to NAEP assessment results.
- The Test and Honors File – Provides a list of honors and standardized test results that were included on the transcripts.
- The Transcript File – Provides a complete list of all courses appearing on the transcripts of students in the study.
- The SD/LEP File – Provides detailed information on students with disabilities and/or limited English proficiency.

Four additional NAEP assessment files contain proficiency estimates (also described as plausible values, as discussed in Chapter 7) for each student who completed NAEP. These are:

- The 1998 NAEP 25-Minute Writing Assessment Data File;
- The 1998 NAEP 50-Minute Writing Assessment Data File;¹⁰
- The 1998 NAEP Reading Assessment Data File; and
- The 1998 NAEP Civics Assessment Data File.

These files contain NAEP scores for the total number of 1998 graduates who participated in both the specific NAEP assessment and the transcript study. However, students who did not meet the graduation requirements were later excluded from the transcript study. Their data is present only in the NAEP assessment files and not in the transcript data files.

This report provides a brief description of the sampling of schools and students (Chapters 2 and 3), the data collection procedures (Chapter 4), data processing procedures (Chapter 5), and the weighting procedures (Chapter 6). Chapter 7 describes the codebooks and data files that are included in this report, which can be found in Appendices D through N. Appendices A-C contain the questionnaires used in this study.

¹⁰ The 50-minute writing assessment file does not contain proficiency assessments. Instead, a categorical determination was assigned for the assessment.

2. BACKGROUND: SAMPLE DESIGN

The 1998 High School Transcript Study (HSTS) was designed to allow an analysis of the coursetaking patterns of students who graduated from American public and nonpublic high schools in 1998. It was further designed so that data on students' coursetaking patterns can be linked to the 1998 National Assessment of Educational Progress (NAEP) assessment results. Since studies similar to the 1998 HSTS were conducted on 1982, 1987, 1990, and 1994 graduates, changes in these patterns and relationships to NAEP performance in these years can also be studied.¹

The HSTS used a subsample of primary sampling units (PSUs) and schools from the 1998 NAEP assessment for grade 12 students. The HSTS used the NAEP target sample of students in these subsampled schools. This chapter describes aspects of the 1998 NAEP sample design that affect the transcript study. Chapter 3 describes aspects of the selection of PSUs, schools, and students that are specific to the transcript study.

2.1 1998 NAEP Sample Design

The 1998 National Assessment of Educational Progress used a multistage probability sample design. Counties or groups of counties were the first-stage sampling units, and elementary and secondary schools were the second-stage units. The third stage of sampling consisted of the assignment of sessions by type to sampled schools and the assignment of sample types to sampled schools. The session type refers to the subject(s) being assessed, while the sample type refers to the specific criteria for inclusion that were applied to the session (see Section 2.4 for a discussion of the inclusion criteria). The fourth stage involved selection of students within schools and their assignment to session types.

A total of 94 Primary Sampling Units (PSUs) were included in the sample, and a sample of 733 schools actually participated in the assessment for grade 4, 761 schools for grade 8, and 608 schools for grade 12. Various blocks or packages of exercises were administered to students in these schools.

¹ The 1987, 1990 and 1994 transcript data were collected by Westat in coordination with the 1987, 1990, and 1994 NAEP (Thorne et al., 1989; Legum et al., 1993; Legum et al., 1997). The 1982 data were collected by the National Opinion Research Center as part of the High School and Beyond project (Jones et al., 1983a).

2.2 Selection of NAEP Primary Sampling Units

In the first stage of sampling, the United States (the 50 states and the District of Columbia) was divided into geographic primary sampling units. Each PSU met a minimum size requirement (a 1990 census population of at least 60,000 in the Northeast and South and 45,000 in the Midwest or West regions) and comprised a metropolitan statistical area (MSA), a single county, or (more often in the case of nonMSA PSUs) a group of contiguous counties. In the case of New England MSAs, which are not formed from whole counties, the corresponding New England County Metropolitan Areas, which are defined in terms of whole counties, were designated as the PSUs. Each PSU was contained entirely within one of the four geographic regions defined in Table 2-1. Each region contains about one-fourth of the U.S. population. These regions were used to stratify the sample of PSUs, ensuring that each region was adequately represented in the various assessment samples.

In a few cases, a metropolitan statistical area crossed region boundaries. Such MSAs were split into two or more PSUs as necessary. For example, the Cincinnati OH-KY-IN MSA was split into the Cincinnati OH-IN PSU in the Central region and the Cincinnati KY PSU in the Southeast region.

Table 2-1. NAEP geographic regions used for stratification

Northeast	South	Midwest	West
Connecticut	Alabama	Illinois	Alaska
Delaware	Arkansas	Indiana	Arizona
District of Columbia	Florida	Iowa	California
Maine	Georgia	Kansas	Colorado
Maryland	Kentucky	Michigan	Hawaii
Massachusetts	Louisiana	Minnesota	Idaho
New Hampshire	Mississippi	Missouri	Montana
New Jersey	North Carolina	Nebraska	Nevada
New York	South Carolina	North Dakota	New Mexico
Pennsylvania	Tennessee	Ohio	Oklahoma
Rhode Island	Virginia*	South Dakota	Oregon
Vermont	West Virginia	Wisconsin	Texas
Virginia*			Utah
			Washington
			Wyoming

* That part of Virginia that is part of the Washington, DC-MD-VA metropolitan area is included in the Northeast region; the remainder of the state is included in the Southeast.

The 22 largest PSUs in the United States were included with certainty (that is, with probability = 1). The remaining smaller PSUs were not guaranteed to be selected for the sample (that is, they were included with probability < 1). These were grouped into a number of noncertainty strata and one PSU was selected from each stratum. Within each major stratum or subuniverse, further stratification was achieved by ordering the noncertainty PSUs according to several additional socioeconomic characteristics, yielding 72 strata.

The strata were defined so that the aggregate of the measures of size of the PSUs in a stratum was approximately equal for each stratum. The size measure used was the population from the 1990 Census. The characteristics used to define strata were the percentage minority population, percentage change in total population since 1980, per capita income, percent of persons age 25 or over with college degrees, percent of persons age 25 or over who completed high school, and the civilian unemployment rate. Up to four of these characteristics were used in one subuniverse. For each subuniverse, the characteristics used were chosen by modeling PSU-level mean reading proficiency scores for 1988, 1990, and 1992. Then one PSU was selected with probability proportional to size from each of the 72 noncertainty strata. That is, within each stratum, a PSU's probability of being selected was proportional to its population.

The final sample of 94 PSUs was drawn from a population of about 1,000 PSUs. Primarily because of the use of MSAs as PSUs (they varied greatly in size), PSUs varied considerably as to their probability of selection. In each region, noncertainty PSUs were classified as metropolitan (MSA) or nonmetropolitan (nonMSA). The 36 selected noncertainty MSA PSUs had probabilities ranging from 0.03 to 0.56, while the 36 nonMSA PSUs had probabilities ranging from 0.03 to 0.10. Parts of 44 states were included in the main sample PSUs. Ninety-four PSUs were selected for the main NAEP sample (22 certainty and 72 noncertainty). A subset of 58 of these same PSUs was randomly selected for the HSTS sample. The major strata, or subuniverses of noncertainty PSUs, are shown in Table 2-2.

2.3 Selection of NAEP Schools

For NAEP, after the PSUs were selected, the next step was to select the schools within the PSUs. For the second stage of sampling, a frame list of 12th grade schools was formed within each PSU. There were 4,513 public and 4,853 nonpublic schools on the final second-stage sampling frame.

Table 2-2. Noncertainty PSU strata

Region	Number of strata for MSA PSUs	Number of strata for nonMSA PSUs	Total
Northeast	6	4	10
South	12	12	24
Midwest	8	12	20
West	10	8	18
Total	36	36	72

The public schools (including Bureau of Indian Affairs [BIA] schools and Department of Defense Education Activity [DoDEA] schools) and nonpublic schools (including Catholic schools) in each PSU were listed. The lists of schools were obtained from two sources. A list of public, BIA, and DoDEA schools, which is maintained by Quality Education Data, Inc. (QED),² and included information obtained from the 1994-95 NCES Common Core of Data (CCD), was obtained in early March 1997. Regular public schools are schools with students who are classified as being in a specific grade (as opposed to schools having only “ungraded” classrooms). These include statewide magnet schools and charter schools.

Lists of Catholic and other nonpublic schools were obtained from the Private School Survey (PSS) conducted by the National Center for Education Statistics. The PSS list of schools is an ongoing registry of private schools. The registry is updated prior to the survey through two sources. The first source, called the list frame, is a conglomeration of a number of lists from several associations, states, and so on. Although the list frame attempts to provide complete coverage of the private school universe, it needs to be supplemented with a second source. The second source uses an area frame to identify and represent schools not on the list frame. The area samples are conducted first by randomly selecting primary sampling units (PSUs); these are single counties or groups of counties from the area frame, which consists of all counties in the nation. Within each selected PSU, a complete list of schools is gathered from a variety of means, and schools not on the list frame are identified and added to the list frame of nonpublic schools. The probabilities of selection for schools on the PSS list ranged from 0.06 to 1.00

² Quality Education Data, Inc. (Denver, CO) (QED) is a privately maintained database of public and private schools in the United States that provides an annual listing of all schools and school districts in the United States, released in November of each year. This listing corresponds to the previous school year. It includes information about each school's name, mailing address, location address, district name, FIPS state number, Office of Education district number, number of students, number of teachers, grades served, and other sociodemographic data.

(most were equal to 1.00). A weight component was computed so that these selected PSS nonpublic schools represented themselves and also represented the non-PSS nonpublic schools for non-PSS PSUs.

For each school in the frame, estimates were made of the number of eligible students. The QED and PSS files give total enrollment and the grade range for each school, thus providing the average enrollment per grade. Schools were selected across all PSUs, systematically from a sorted list with probabilities proportional to assigned measure of size, which was a function of the average enrollment per grade. The sorting variables included NAEP region, private/public classification, type of location, high/low minority group, PSU stratum, and grade enrollment. To increase cost efficiency in sampling, samples were designed to include more nonpublic and high-minority public schools, and more relatively large schools.

Each public school that was considered high minority (i.e., with over 15 percent black and/or Hispanic enrollment) was given double the probability of selection of a public school which was not considered high minority and which was of a similar size, in the same PSU. Such high-minority schools were oversampled to enlarge the sample of black and Hispanic students, thereby enhancing the reliability of estimates for these groups. Given a specified sample size, this procedure reduces slightly the reliability of estimates for all students as a whole and for those not black or Hispanic.

In NAEP, each private school was given triple the probability of selection of a low-minority public school of similar size from the same PSU. These greater probabilities of selection were used to ensure adequate samples of private school students in order to allow the derivation of reliable NAEP estimates for such students. In HSTS, however, the oversampling of private schools was reversed by taking a private school subsample from the NAEP sampled schools at only one-third the sampling rate of the corresponding public school sample (see Chapter 3).

The QED files do not contain schools that opened between 1996 and the assessment dates. Therefore, special procedures were implemented to be sure that the NAEP assessment represented students in new public schools. Small school districts—those that contained only one eligible school—were handled differently from large school districts, which contained more than one eligible school. In small school districts, the schools selected were thought to contain all students in the district that were eligible for the assessment. Districts containing these schools were asked if other schools with 12th grade existed and, if so, they were automatically included in the assessment.

For large school districts a district-level frame was constructed from the schools on the QED file. Then districts were sampled systematically with probabilities proportional to a measure of size. In most cases, the measure of size was total district enrollment, but in very small districts a minimum measure of size was used. Each sampled district was asked to update the list of eligible schools derived from information on the QED files. Frames of eligible new schools were then constructed for 12th grade, and samples of new schools were selected systematically with probability proportional to eligible enrollment using the same sampling rates as for the old schools. As a result of this process, three new public schools were selected.

Potential substitute schools were selected for all sampled schools in the 1998 NAEP where a close match could be identified. An attempt was made to preselect (before field processes began) a maximum of two substitute schools for each sampled public school (one in-district and one out-of-district) and each sampled Catholic school, and one for each sampled nonCatholic nonpublic school. A nonparticipating school was replaced by a substitute when the participating school was considered a final refusal. To minimize bias, a substitute school resembled the original selection as much as possible.

Substitutes were assigned by matching approximately on the following attributes:

- Affiliation (public or private);
- Estimated number of eligible students; and
- Minority composition.

A substitute was always selected from the same PSU as the refusing school. When school nonparticipation was due to district refusal, none of the schools in the refusing district were considered substitute candidates. However, when substituting for school-level (rather than district-level) refusals, preference was given to substitute candidates in the same district. Of the 608 participating grade 12 sampled schools, 48 were substitutes.

2.4 Assignment of Sessions to Schools for NAEP

Three subjects were assessed in different types of assessment sessions. The assessment subjects were writing, civics (writing and civics were combined into one session, as the directions and timing of the sections were the same), reading (the reading assessments, at grades 8 and 12, included

some booklets that consisted of two 25-minute blocks of questions and others that consisted of one 50-minute block, but were combined in one session type), and civics trend. The last time that civics was assessed was in 1988, and since that time the civics frameworks and items for NAEP have changed. In order to measure trends with the past, yet also measure students' knowledge in relation to the new frameworks, two different civics assessments were conducted in 1998. Civics trend used the identical items and procedures from the 1988 assessment, while the new frameworks were evaluated with new civics items (that were field-tested in 1997).

Each 12th grade was allocated a number of sessions, based on the estimated number of grade-eligible students, as shown in Table 2-3.

Table 2-3. Allocation of sessions

Estimated number of grade eligibles	Number of sessions allocated
1 – 30	1
31 – 60	2
61 – 90	3
91 – 120	4
121 or more	5

The sessions were allocated to 12th grade sampled schools by placing schools in the order used for sampling and allocating the appropriate number of sessions from the following repeated sequence (W denotes writing/civics, R denotes reading and C denotes civics trend): R, W, W, R, W, W, R, W, W, R, W, W, C, W, W, R, W, W, R, W, W, W, R, W, W, W, R, W, W, R, W, W, R, W, W, C, W, W, R, W, W, R, W, W, W, R, W, W, W. The sequence contains 34 W's, 13 R's, and 2 C's. It was designed to ensure the maximum feasible spread of assessment types among schools, while ensuring that close to 69 percent of the selected students were assigned to writing/civics, 27 percent to reading, and 4 percent to civics trend, as summarized in Table 2-4.

Table 2-4. Allocation of sessions to schools

Session name	Number of sessions	Percent of selected students assigned to session
Writing/civics (W)	34	69
Reading (R)	13	27
Civics Trend (C)	2	4

Schools with 31 or more eligible students were always assigned at least one writing/civics session. Schools with 91 or more eligibles were almost always assigned a minimum of one reading session. Many schools were awarded multiple sessions of the same type or multiple sessions of different types. This did not necessarily mean that the school had to physically conduct multiple sessions of a given assessment type, but the assignment of session types determined the proportions of selected students within the school that were assigned to each type.

In order to determine the effect of using different criteria for excluding students from the assessment, two different sample types (S2 and S3) were assigned to the session types assigned to schools. In sample type 2 schools, the 1996 exclusion criteria were used. In sample type 3 schools, the 1996 exclusion criteria were used and accommodations were offered to students with disabilities (SD) and students with limited English proficiency (LEP). For schools assigned a reading session, sample type was assigned to schools separately so that 50 percent of the schools assigned reading were assigned sample type 2 and 50 percent were assigned sample type 3. The schools were placed in the order of sampling, then sample types were assigned to schools with a reading session by alternating sample types 2 and 3. Sample type was assigned so that a variety of schools with respect to region, school type, urbanization, and size were in each sample type. For writing/civics sessions, only sample type S3 was assigned. For civics trend sessions, only sample type S2 was assigned.

2.5 Sampling Students

In the fourth stage of sampling, the sample of students within sampled schools was systematically drawn from school-prepared lists of eligible students. Student Listing Forms (SLF) were prepared for each participating school; all enrolled students of the 12th grade were to be entered on the SLFs. Student samples also included oversampling of black and Hispanic students in schools with low minority enrollment, and oversampling SD/LEP students in public schools assigned to reading, and were specified through the use of Session Assignment Forms (SAF).

Up-to-date information on grade enrollment was obtained for sampled schools through two field processes. Scheduling assessment dates with schools and being on site at the school at the time of sampling and the assessment allowed field staff to obtain updated information on the number of grade eligibles.

The district supervisor generally carried out the sampling of students a week before the assessment. Student Listing Forms were prepared in each participating school. All enrolled 12th grade students were to be entered on the SLF in any order convenient to the school, or the school could produce a computer-generated list. Before carrying out the sampling, the district supervisor reviewed the form and made comparisons with other enrollment information to ensure that the list included all eligible students. Once the list was determined to be complete, a sequential line number was assigned to each student.

The sample of students to be selected in each school was derived in the following manner. A maximum sample size of 150 students was set for each school. In schools that, according to information on the frame, had fewer than 150 eligible students, each eligible student enrolled at the school was selected in the sample for one of the sessions assigned to the school. In the larger schools, a sample of students was drawn and students were assigned to sessions as appropriate.

The assignment of students to sessions was completed in the following way. After the students were numbered on the Student Listing Form, the field worker referred to the school's designated SAF. There, the line numbers for each of the school's assigned sessions were listed. For instance, a Civics Trend session might include the students listed on lines 4, 9, 14, 19, 24, 30, and so on, with a different sequence of line numbers for the students designated for a Reading session.

The field workers for the 1998 High School Transcript Study were drawn from the pool of NAEP field supervisors. To avoid confusion, the data collection personnel for the HSTS are referred to simply as field workers. If the field worker found that the line numbers, when applied to the numbered list of eligible students assembled in the field for each school, generated a sample in excess of 170 students, he or she called a field supervisor. New line numbers based on the actual number of eligible students were generated on a personal computer and relayed to the field worker. A similar revision to the line numbers was made in a school with a sampling interval in excess of 1.0 and eligible enrollment less than 80 percent of that initially estimated. In this case, the sample size was increased to the appropriate level. This procedure provided a suitable compromise between control over the sampling rate within each school and operational autonomy and flexibility for field workers.

In all cases where new line numbers were generated, sampling intervals were sent to Westat's central office and stored for use in sample weighting. Field workers were not required to derive or record within-school sampling rates.

Students were assigned to the sessions systematically, in proportion to the number of sessions of each type allocated to the school. To control the student sampling operations as closely as possible, Westat generated a Session Assignment Form for each school where sampling was to be carried out. This computer-generated form specified:

- The types of sessions that were to be administered at the school;
- The line numbers (from the SLF) specifying the students to be drawn into the sample;
- The minimum and maximum number of students listed on the SLF that could be accepted without requiring revision to the within-school sampling rates;
- Whether accommodations were to be offered to SD/LEP students;
- Instructions and line numbers for oversampling black and Hispanic students in public schools with low minority enrollment and SD/LEP students in schools assigned reading; and
- Special instructions as appropriate for the 1998 SD/LEP Questionnaire.

It became necessary, because of updated grade enrollment numbers, to revise the session allocation structure for some smaller-than-expected schools with more than one session type initially assigned. Smaller-than-expected schools were defined as having a potential of fewer than 12 students assigned to a particular session type. For example, if two writing/civics and one reading session were assigned, and the number of grade eligibles was updated to 30 students, then only 10 would be assessed in reading.

In this case, and in general, for smaller-than-expected schools where the number of grade-eligibles per session type assigned (without regard to the number of sessions assigned for each type) was 12 or more (15 in the example), all session types were kept and students were split evenly across the session types. Thus, in the example given here, 15 students would be assigned to reading and 15 to writing, rather than the initial sample allocation number of 10 and 20, respectively. If the number of grade-eligibles per session type assigned was fewer than 12, just one session type was kept at random, and a weight adjustment factor was computed as the ratio of the number of sessions assigned to the number of sessions assigned for the session type that was kept. This weight adjustment accounts for the dropping of one or more session types.

In public schools with low minority enrollment, an oversample of black and Hispanic students was selected. After the initial sample was selected, the nonselected black and Hispanic students

were identified and listed. These students were sampled to a total that was expected to be the same number of black and Hispanic students already selected. In practice, however, if the number of nonselected students was less than the number of selected students, then all nonselected black and Hispanic students were also to be assessed. Otherwise, these students were sampled so that their overall within-school probability of selection was twice the rate of other students.

Line numbers were generated to split the additional sample of black and Hispanic students into sessions using the session allocation rates applied to the initial sampling procedure. Thus, if the school was assigned two sessions of writing/civics and one of civics trend, two-thirds of these extra black and Hispanic students were assigned to writing/civics, and one-third to civics trend.

The procedures for assessing students with disabilities and limited English proficient students (SD/LEP) varied by sample type. Those in sample type 3 were offered accommodations not available to other students or to SD/LEP students in sample type 2. Oversampling procedures were applied to SD/LEP students as a measure to ensure an adequate sample size from both sample types 2 and 3 for reading. In this way, comparisons of the effect of offering accommodations to students have enhanced power to detect effects.

The general intent of oversampling within each school assigned at least one reading session was to select SD/LEP students at twice the rate at which non-SD/LEP students were sampled (or to include all SD/LEP students if there were not sufficient numbers to permit sampling at twice the rate). There was no oversampling of schools as part of this procedure. In each school oversampled for SD/LEP students, the initial desired sample of students was drawn for each session assigned, from the full list of eligible students. As previously stated, black and Hispanic students were oversampled in public schools in low-minority areas. Among students not selected for either of the two prior sampling operations for this school, the SD/LEP students were identified. A sample from among these was drawn, using a sampling rate that would achieve the double sampling rate required overall. Again, the weighting procedures ensured that the results were not biased as a result of the relative under-representation of SD/LEP students from smaller schools and relative under-representation of black and Hispanic students from smaller low-minority schools.

2.6 Students not Included in the Assessment

Once the sample of students was selected, school staff were asked to identify any students with a disability and any students classified as limited English proficient. The SD/LEP Questionnaire was then distributed to the school staff member most knowledgeable about the student, as described in Section 4.5. The questionnaire collected information about the student's disability/language proficiency and any special services provided by the school.

School staff were also asked to determine whether any of the students identified as SD or LEP could not participate meaningfully in the assessment. These students were not invited to the assessment and were coded as "excluded" to distinguish them from absent students. Transcripts for these students are, however, included in the transcript study.

3. SELECTION OF PRIMARY SAMPLING UNITS, SCHOOLS, AND STUDENTS FOR THE 1998 HIGH SCHOOL TRANSCRIPT STUDY

The purposes of the 1998 High School Transcript Study (HSTS) were to gather data on a nationally representative sample of students who graduated from American public and nonpublic high schools in 1998 and to gather data that could be linked to NAEP results. For the HSTS sample of students to be as representative as possible, it included a subsample of NAEP PSUs, and subsequently subsampled schools with 12th grades within the PSUs that were selected for NAEP, regardless of whether they participated in NAEP. A representative sample of students was included from each school. When possible, the students selected for the transcript study were the same as those selected for NAEP. When this was not possible, a systematic sample of students was drawn from the school. The PSU sample, the school sample, and the student sample are described in detail in the following sections.

3.1 PSU Sample

As discussed in Chapter 2, the 1998 NAEP sample included the selection of PSUs as the first stage of sampling. To obtain a substantially smaller number of schools (322 12th grade schools were selected from the 852 originally sampled in NAEP) in order to reduce field costs, a subsample of the NAEP PSUs was selected for the HSTS. All 22 certainty PSUs and half of the 72 noncertainty PSUs were selected. For selecting the 36 noncertainty PSUs, the 72 NAEP sampled PSUs were combined into pseudostrata based on region, urbanicity, and socioeconomic characteristics. Then the PSUs were selected based on the following probabilities of selection:

- A probability of selection equal to one was assigned to the certainty PSUs and
- A probability of selection equal to one-half was assigned to the noncertainty PSUs.

One PSU was selected randomly within each noncertainty pseudostratum. A total of 58 PSUs were selected for the HSTS.

3.2 School Sample

The 1998 HSTS sample comprised only schools selected for the NAEP main sample that had 12th-grade classes and were within the 58 PSUs selected for the HSTS. There were 606 eligible schools that satisfied this criterion, of which 366 were public and 240 were nonpublic. In the next step of selection, a subsample of 322 schools was selected, consisting of 269 public schools and 53 nonpublic schools. To create the subsample of schools for the HSTS, the following probabilities of selection were assigned to offset the increased probability of selection for nonpublic schools (three times that of public schools with low numbers of minority students) that occurred in the NAEP sample:

- A probability of 1/2 was assigned to public schools in certainty PSUs;
- A probability of 1/6 was assigned to nonpublic schools in certainty PSUs;
- A probability of 1 was assigned to public schools in noncertainty PSUs; and
- A probability of 1/3 was assigned to nonpublic schools in noncertainty PSUs.

Prior to sampling, the schools were sorted in the sort order of the original sample procedure relating to schools in NAEP (refer to Section 2.3 for the sorting variables). An oversample of nonpublic schools was considered important for the NAEP sample but was not considered desirable for the HSTS sample. Because nonpublic schools tend to be smaller than public schools, the collection cost per transcript is higher.

In order to maintain as many links as possible with NAEP scores, where schools refusing to participate in NAEP were replaced by substitute schools, the substitute schools, not the refusals, were asked to participate in the HSTS. Of the 322 schools in the original sample, 264 original/substitutes participated in the HSTS survey, of which 241 were originally sampled. Of the 264 participating schools, 232 schools cooperated with both HSTS and NAEP and the links for the students were maintained, 10 schools cooperated with HSTS and NAEP but the links for the students were not maintained, and 22 schools cooperated with HSTS but not with NAEP.

3.3 Student Sample

For schools participating in both NAEP and HSTS, the same students were included in the two samples where possible. For privacy reasons, the only means of identifying the students participating in NAEP was a list left in the school office. Since the NAEP assessments were administered from January through April 1998, the schools were asked to retain the NAEP administration schedules until the HSTS data collection in the summer and fall of 1998.¹ The administration schedules are forms produced specifically for each school. They include the assessment booklet ID's that are assigned to each school, which are listed sequentially on the administration schedules. Once the student sample is drawn, the selected student's name is recorded on the administration schedule for the type of session for which he or she was selected. As this is done, the booklet ID on that line becomes the student's NAEP ID number. This is the only place where selected students' names are recorded. Other demographic information is also recorded on the administration schedule, which is shown in Exhibit 3-1.

For schools that participated in NAEP but were missing their administration schedules, and for schools that agreed to provide transcripts but did not participate in the NAEP assessment, the field workers sampled the students using the following rules:

- If 60 or fewer students were in the senior class, all students were selected for the study.
- If more than 60 students were in the senior class, the field worker drew a systematic random sample of 50 students.

To draw a sample, the field worker obtained a complete list of students in the senior class, numbered each student sequentially, and then entered the number of students in the class and the number of transcripts needed (50) onto a sampling form. After determining the number of students in the senior class, the field worker calculated a sampling interval. A random start was drawn from a supplied list of random numbers, and a systematic sample was drawn based on the random start and the sampling interval. The field worker then wrote the names of the sampled students on a Transcript Request Form (TRF) and gave it to the school staff to draw the transcripts. The TRF also provided a place to record the

¹ NAEP asked schools to retain the administration schedules until the end of the calendar year in case it became necessary to use them to resolve ID-related questions. For reasons of confidentiality, the schools that were not in the transcript study were requested to destroy these materials by June 30, 1998.

[illegible]

students' graduation status, gender, race/ethnicity, birth month, birth year, disability status, LEP status, receipt of Title I services, and National School Lunch Program participation. To maintain confidentiality, the field worker removed the students' names before returning a copy of the TRF to Westat along with the transcripts.

When field workers went to the schools to collect the transcript data, they had been supplied with sets of labels for each student NAEP ID at the school. They had also received a Transcript Request Form produced for each school, with each ID listed on a line of the form, along with the demographic information that had been collected on the student at the time of the assessment. As they collected the transcripts, they attached the ID labels to them to identify the student to whom they belonged. At the same time, they made sure that any other identifying information was erased or obscured, so that the student could not be identified. For schools that had not participated in NAEP, a set of labels was created with newly assigned ID numbers for the students selected in that school. In those schools, the TRF was produced with the new ID numbers, but with space to record all of the demographic information that was collected.

A total of 28,764 students were selected for inclusion in the HSTS. Of these, 27,183 students were from schools that maintained their NAEP administration schedules and were identified by their NAEP booklet numbers. Another 500 students were from schools that participated in NAEP but had lost the link between student names and NAEP booklet numbers, and 1,081 were from schools that did not participate in NAEP.

Table 3-1 displays the number of eligible schools in the sample and the number and percent of schools from which we collected transcripts, by linking category. Where it is indicated that transcripts were collected, it means they were usable transcripts of graduating students.

Table 3-1. Response rates of eligible schools by linking category, unweighted

School participation status	Number of schools in sample	Number of schools where transcript data were collected	Percent of schools where transcript data were collected
Original school participated in NAEP—IDs linked to NAEP IDs	221	211	95.5
Original school participated in NAEP—IDs not linked to NAEP IDs	8	8	100.0
Original school did not participate in NAEP	72	22	30.6
Eligible original sampled schools	301	241	80.0
Substitute school participated in NAEP—IDs linked to NAEP IDs	21	21	100.0
Substitute school participated in NAEP—IDs not linked to NAEP IDs	2	2	100.0
Total substitute schools	23	23	100.0
Total original and substitute schools	324	264	81.5

Table 3-2 displays the number of sampled students in the participating (original and substitute) schools and the number and percent of completed transcripts of graduates that were processed.

Table 3-2. Percent of sampled students who were graduates and for whom completed transcripts were received

School participation status	Number of students in sample	Number and percent of sampled students who were graduates and for whom completed transcripts were received*
School participated in NAEP—IDs linked to NAEP IDs	27,183	22,804 86.4
School participated in NAEP—IDs not linked to NAEP IDs	500	461 93.0
School did not participate in NAEP	1,081	953 88.9
Total	28,764	24,218 86.6

* This number reflects the number of usable transcripts collected.

Because sampling was performed in most schools prior to graduation, not all sampled students were, in fact, graduates. Only graduates, however, were eligible for the transcript study. We were able to determine that of the 28,764 students in the sample, 25,248 actually graduated and that 3,328 did not. Of the remaining 188 students, we imputed 174 as graduates and 14 as not. Thus, from the 25,422 graduates Westat collected and processed 24,218 transcripts. That is, Westat was able to obtain 98.0 percent of the transcripts of eligible students. Table 3-3 displays the response rates for graduates in the eligible participating schools.

Table 3-3. Response rates of graduates, unweighted

School participation status	Known graduates	Imputed graduates	Known and imputed graduates	Number of transcripts of known graduates collected	Percent of transcripts of known graduates collected	Percent of transcripts of known and imputed graduates collected
School participated in NAEP—IDs linked to NAEP IDs	23,803	97	23,900	22,804	98.6	98.2
School participated in NAEP—IDs not linked to NAEP IDs	477	0	477	461	97.5	97.5
School did not participate in NAEP	968	77	1,045	953	99.3	92.0
Total	25,248	174	25,422	24,218	98.6	98.0

Table 3-4 displays the weighted response rates for NAEP, the transcript study, and the linked schools.

Table 3-4. Response rates for NAEP, transcript study, and linked schools, weighted

	Weighted school response rate before substitution (in percent)	Weighted school response rate after substitution (in percent)	Weighted student response rate (in percent)	Overall response rate (in percent)
Overall NAEP	75.2	81.6	79.6	65.0
25-Minute Writing	69.7	78.0	79.7	62.1
50-Minute Writing	69.7	78.0	80.4	62.7
Civics	69.7	78.0	79.4	61.9
Reading	69.7	78.2	80.1	62.6
Transcript Study	76.6	85.3	98.3	83.8
Linked Schools				
25-Minute Writing	65.7	73.5	81.6	60.0
50-Minute Writing	65.7	73.5	82.4	60.6
Civics	65.7	73.5	80.7	59.3
Reading	65.6	73.2	82.5*	60.4

* The student response rate reflects all students within S2 and S3 schools. Therefore, some students that are in the response rate calculation are not in the reading reporting population, as defined in Section 6.5.7. Section 6.5.7 discusses how a portion of students in S3 schools assigned reading is used to comprise the reporting population for the NAEP reading assessment.

4. DATA COLLECTION PROCEDURES

4.1 Training NAEP Field Supervisors as Data Collectors

The field workers for the 1998 High School Transcript Study were drawn from the pool of NAEP field supervisors. They were trained in the data collection procedures for HSTS in December 1997. This training was conducted by the HSTS Curriculum Specialist/Coding Supervisor and took place over one full day. The training consisted of three sessions. The purpose of the first session was to establish the background knowledge needed to help field workers to make informed decisions when collecting information in the schools, and to explain why attention to detail and accuracy would be crucial in ensuring the quality of HSTS data. The second training session was held to familiarize field workers with the HSTS materials and forms and with the variety of materials they could expect to find in the schools. The third session provided an opportunity for field workers to work with sample catalogs and transcripts, and to fill out practice forms, as they would do using the actual materials for the HSTS.

The first training session consisted of a presentation describing the purposes of the HSTS, the procedures Westat uses in handling and processing HSTS data, and the best sources of data to obtain from schools to provide Westat with the needed data. Specific examples were used throughout the presentation.

During the second session, field supervisors were shown examples of various types of high school records and materials, including school- and district-level catalogs, course lists, transcripts, and all the forms used and completed for the HSTS. The information on each of these materials was cross-referenced to the data needed for the HSTS at the school and student levels. Transparencies of screen prints of the transcript data entry and course coding systems were shown to demonstrate how the information from the specific materials would be entered.

The third training session consisted of completing sets of exercises, designed to provide the field workers with hands-on experience in examining school materials and filling out the forms they would use. The practice materials consisted of copies of actual catalogs, course lists, and transcripts obtained in the 1994 HSTS (with all identifying information deleted).

The first set of exercises was completed by the group as a whole, using transparencies of the materials and an overhead projector. The second set was completed in pairs or small groups, and the third set was completed individually and collected for review by supervisory staff. Errors or misconceptions were corrected and discussed with the field workers prior to their leaving the training session. Sample catalogs included a course list, extracts from a large catalog, and a smaller catalog. The sample materials were selected to give field workers a sense of the variety of materials they might expect to find in schools with respect to the amount of information available, the physical layout of the materials, and the ease or difficulty of accessing the information in the materials. Transcripts were examined in this exercise to show a number of ways that special education, for example, might be indicated, as well as indicators for transfer courses, remedial courses, honors courses, off-campus location courses, or courses for students with limited English proficiency.

4.2 Contacts with States, Districts, and Schools

In September 1997, superintendents and principals were notified about the transcript study through the Summary of School Tasks which was included in a mailout to all schools selected for NAEP. This summary included information on several aspects of the main NAEP study, as well as the notification of the transcript study. In December 1997, district superintendents of participating 12th-grade schools sampled for the main NAEP and selected for the HSTS were mailed additional information concerning the HSTS. Items in the package included the following:

- An informational letter to school superintendents from the Project Officer of NCES (Exhibit 4-1);
- A list of schools in the district selected for the 1998 HSTS; and
- A summary of school transcript activities (Exhibit 4-2).

For contacts with school-level personnel, field workers were provided with the following materials:

- An informational letter to principals from the Project Officer of NCES (Exhibit 4-3).
- The summary of school transcript activities.

Exhibit 4-1. Superintendent's letter from Project Officer



U. S. DEPARTMENT OF EDUCATION
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT

NATIONAL CENTER FOR EDUCATION STATISTICS

Dear Superintendent:

As described in previous mailings to your district, the 1998 High School Transcript Study is being conducted in conjunction with the 1998 National Assessment of Educational Progress (NAEP). The purpose of this study is to supply data to educational researchers and policy analysts on course-taking patterns and to examine the relationship of these patterns to achievement in secondary schools sampled in the 1998 NAEP. NAEP schools are included in the sample in order that NAEP data and transcript data can be linked. The participation of all selected schools is needed to make the results of the transcript study comprehensive, accurate, and timely.

A list of the NAEP schools in your district selected for this study is enclosed. Detailed information on transcript activities in the school accompanies this letter. No student time is involved; students' names and other individually identifying information will be removed from copies of the transcripts before they leave the school; schools will be reimbursed at the standard rate for supplying transcripts.

The activities for Phase I will be conducted at the same time NAEP supervisors are in the schools selecting the NAEP sample. In the summer or fall of 1998, at a time that the schools have indicated are convenient, supervisors will return to the schools to collect the requested transcripts.

The granting of Education Department authority for collection of the transcript data has been made pursuant to the provisions of the Family Education Rights and Privacy Act (FERPA) (20 U.S.C. 1232g) as implemented by 34 CFR 99.31 (a)(3)(ii) and 99.35. These laws and regulations permit an educational agency to disclose records to authorized representatives of the Secretary of Education without the prior consent of the survey participants in connection with the audit and evaluation of Federal and State supported education programs. The privacy of the information schools are asked to supply to the NAEP contractors will be protected as required by FERPA and will be further protected by the removal of names and other identifying information. A copy of the relevant section of FERPA regulations is reproduced on the reverse side of this page.

I would appreciate your cooperation in this important component of the 1998 NAEP. If you have any questions about the study or its procedures, I may be contacted at the Department of Education or you may contact Mark Waksberg of Westat at 1-800-283-6237.

Sincerely,

Steve Gorman
Project Officer

Exhibit 4-2. Summary of school transcript activities

1998 HIGH SCHOOL TRANSCRIPT STUDY

SUMMARY OF SCHOOL ACTIVITIES

This sheet summarizes the High School Transcript Study activities that will be undertaken in 1998. Hopefully, it will provide answers to some of the questions you may have. NAEP supervisors will provide you with a more detailed description of these tasks during telephone and in-person visits to the school.

KEY ASPECTS OF THE HIGH SCHOOL TRANSCRIPT STUDY

- NO STUDENT TIME IS INVOLVED. NAEP staff will work with your school and do as much of the work as possible to minimize the burden.
- Students' names and other individually identifying information will be removed from copies of the transcripts before they leave the school.
- Your school will be reimbursed at your usual rate for providing transcripts.

ACTIVITIES INVOLVING SCHOOLS

Phase 1: January – March 1998

1. The 1998 High School Transcript Study sample will be identified by the NAEP supervisor.
2. Course lists or catalogs will also be requested. Course catalogs will be requested for the following years: 1997-98, 1996-97, 1995-96, and 1994-95.
3. A sample of three transcripts will be requested. One should include regular courses, one special education courses, and one honors courses.
4. The NAEP supervisor will need to review transcripts and course catalogs and collect additional information before leaving your school so that questions about either may be clarified.

Phase 2: Summer – Fall 1998

1. In the Summer or Fall of 1998, NAEP staff will return to your school to collect the requested transcripts of students who graduated.

Exhibit 4-3. Informational letter to principals from Project Officer



— U. S. DEPARTMENT OF EDUCATION
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT

NATIONAL CENTER FOR EDUCATION STATISTICS

Dear Principal:

In conjunction with the 1998 National Assessment of Educational Progress (NAEP), the National Center for Education Statistics, U.S. Department of Education has authorized Westat, the NAEP contractor, to obtain student transcript data from a national sample of secondary schools sampled for the 1998 NAEP. The purpose of the 1998 High School Transcript Study is to supply data to educational researchers and policy analysts on course-taking patterns and the relationship of these patterns to student achievement in secondary schools across the nation.

Your school has been selected to participate in this important study and an informational letter has been sent to your District Superintendent. Your school's participation is needed to make the results of this study comprehensive, accurate, and timely. No student time is involved and schools will be reimbursed at their standard rate for supplying transcripts. Detailed information on the transcript activities and the timeframe for data collection accompanies this letter.

The granting of Education Department authority for collection of the transcript data has been made pursuant to the provisions of the Family Education Rights and Privacy Act (FERPA) (20 U.S.C. 1232g), as implemented by 34 CFR 99.31 (a)(3)(ii) and 99.35. These laws and regulations permit an educational agency to disclose records to authorized representatives of the Secretary of Education without the prior consent of the survey participants, in connection with the audit and evaluation of Federal and State supported education programs. The privacy of the information you are asked to supply to the NAEP contractors will be protected as required by FERPA, and will be further protected by the removal of names and other identifying information. A copy of the relevant section of FERPA regulations is reproduced on the reverse side of this page.

I would appreciate your cooperation in this most important component of the 1998 NAEP. If you have any questions about the study or its procedures, I may be contacted at the Department of Education or you may contact Mark Waksberg of Westat at 1-800-283-6237.

Sincerely,

Steve Gorman
Project Officer

Field workers provided these materials to the school principals and school coordinators during their initial visit to schools to conduct NAEP sampling. They discussed the HSTS with the school coordinator prior to the sampling visit when they called to confirm the sampling date.

Eligible schools participating in NAEP were informed about the 1998 HSTS when they received information about NAEP. Schools were provided with information about participating in the HSTS, including procedures that would be used to ensure confidentiality of the data, and the amount and nature of school staff time required for participating in HSTS. For eligible schools that agreed to cooperate, students sampled for NAEP were included in the HSTS sample, and a brightly-colored Disclosure Notice was placed in their folder by a NAEP field worker or school staff member. This notice, shown in Exhibit 4-4, served two functions:

- It alerted the school personnel that information contained in the student's folder would be used for the HSTS.
- Because of its color, it also served as a visible marker for identifying the folders of students in the HSTS sample to facilitate finding their transcript later.

Once participation in the study was authorized by the district, individual schools were contacted. The contact letter for NAEP, for all schools in which 12th graders were assessed, provided information about the 1998 HSTS. Initial HSTS information requested from schools and collected by NAEP field workers at the time of the NAEP assessment included information which they were asked to provide on the School Information Form (SIF), as well as their school's course catalogs for the four most recent school years, including 1997-98, and three sample transcripts. They were also asked to provide a complete transcript for each graduate in the HSTS sample as soon as graduation information was posted on the transcripts. Information provided on the SIF indicated the appropriate date for the field workers to obtain the transcripts. Schools that were eligible for the 1998 HSTS but that had chosen not to participate in the 1998 NAEP assessment were contacted near the end of the school year.

In originally nonparticipating NAEP schools, notification to the schools included information that the intent was to select a sample of up to 50 students and to provide the same confidentiality safeguards with these samples as with all NAEP students. That is, student names would be removed from any papers that left the school. Field workers also emphasized that a school's participation in the High School Transcript Study would not involve any student time.

DISCLOSURE NOTICE

1998 HIGH SCHOOL TRANSCRIPT STUDY

Date: Spring Quarter 1998
Fall Quarter 1998

A copy of this student's transcript ____ will be ____ has been provided to WESTAT, agent for the U.S. Department of Education, National Center for Education Statistics (NCES). The granting of Education Department authority for collection of the transcript data has been made pursuant to the provisions of the Family Education Rights and Privacy Act (FERPa) (20 U.S.C. 122g), as implemented by 34 CFR 99.31(a)(3)(ii) and 99.35. This disclosure statement fulfills the requirements of provision 34 CFR 99.32 of FERPA.

The High School Transcript Study (HSTS), sponsored by NCES, is being conducted to collect information on current course offerings and course taking in the nation's secondary schools. This student has been selected to participate in HSTS, and data from these records will be combined with others into statistical summaries and tables. No individually identifiable information will be released in any form.

For both NAEP participating and nonparticipating schools, the initial contact by the field worker included a discussion of the following:

- Procedures for obtaining transcripts for the selected students and the method for reimbursing the school for the expense and
- The availability of a course catalog or description.

An appointment was then set to visit the school to prepare the transcript requests and obtain the course catalogs.

4.3 Obtaining Course Catalogs, Sample Transcripts, and Other School-Level Information

Field workers requested sample materials for the HSTS when they first contacted a school and collected these materials when they visited the school for sampling. There were 242 schools that participated in NAEP and also participated in the HSTS (although 10 of these schools did not maintain the NAEP-HSTS links). There were also 22 schools from the original school sample that participated in the HSTS, but did not participate in NAEP. The sample materials included, preferably, a course catalog (a list of courses) offered for each of four consecutive years, from 1994-95 through 1997-98; a completed School Information Form, as shown in Appendix B; and three transcripts of students who graduated in 1998, representing a “regular” student, one with honors courses, and one with special education courses. Since these materials were unique to each school, acquiring them before the collection of the actual transcripts enabled Westat staff to examine them and call a field worker or the school (i.e., before school personnel left for the summer) with any questions that arose during the school year. The catalogs and transcripts collected were also examined by the field worker who filled out a Course Catalog Checklist (Exhibit 4-5) and a Transcript Format Checklist (Exhibit 4-6) for each item collected and sent to Westat.

The field worker also gathered general information about class periods, credits, graduation requirements, and other aspects of school policy. Sometimes this information was documented in the course catalog and at other times in a separate school policy document.

Exhibit 4-5. Course catalog checklist

NAEP School ID: _____

School Name: _____

Supervisor: _____

COURSE CATALOG CHECKLIST

Record each catalog title and check off all items which are identified in the course description materials you have collected.

School Level Materials								
School Year	Catalog Title	Course Title	Course Number	Course Credits	Course Description	Course Level ¹	Special Codes ²	Special Programs ³
1994-95								
1995-96								
1996-97								
1997-98								

District Level Materials								
School Year	Catalog Title	Course Title	Course Number	Course Credits	Course Description	Course Level ¹	Special Codes ²	Where Offered ⁴
1994-95								
1995-96								
1997-98								
1997-98								

¹ Identified as Regular, Honors, A.P, Remedial, Special Education, ESL?

² Does the catalog describe what codes mean?

³ Are Special Programs (Sp. Ed, IB, Vocational, etc.) included in this catalog?

⁴ Does the district catalog identify courses offered at the sampled HSTS school?

Exhibit 4-6. Transcript format checklist

NAEP School ID: _____

Supervisor: _____

TRANSCRIPT FORMAT CHECKLIST

Marked	Not Marked	Not on Transcript	
			1. Student's birthdate
			2. Student's race/ethnicity
			3. Student's gender
			4. Student's LEP/LEP status
			5. Student's graduation date
			6. Years attending this school
			7. Type of diploma awarded
			8. <u>When</u> a course was taken (year and semester)
			9. For a single course:
			a. course name
			b. number of credits awarded
			c. length of course (one year, semester, or other)
			d. grade received
			e. level of course (honors, remedial, SpEd, regular)
			f. transfer credit from another high school
			g. taught in another language (or ESL course)
			h. vocational courses
			i. location, if not taught at this school site
			10. Total number of credits received
			11. "Weighting" of course credits/grades (for honors or remedial levels)
			12. Are abbreviations or codes used on the transcripts? If so, indicate on the back of this form what they are and what they mean for those that are not obvious.

4.3.1 Catalogs

Course catalogs were carefully reviewed at the school. Field workers verified that the catalogs contained all of the courses that 1998 12th graders could have taken in high school, including vocational, remedial, honors, special education, or off-campus courses, or courses taught in a language other than English. If these course listings were not in the catalog, every effort was made to obtain additional information from school personnel to document the existence of such courses and to describe them.

Our prior experience in coding course catalogs for previous HSTS studies led us to identify the following levels of priority for the type of catalog to request:

1. A school-level catalog providing course titles and descriptions;
2. A district-level catalog, if it indicated which courses were offered at the HSTS participating school;
3. A course list by department that included general descriptions of course offerings by department;
4. A school-level course list without descriptions; or
5. A district-level catalog without any indication of which courses were offered in specific schools.

Field workers filled out a Course Catalog Checklist for the catalogs they obtained. This checklist served two purposes:

- It guided field workers in obtaining materials with the maximum amount of information possible that would be useful in the HSTS.
- It provided Westat staff with a quick way to review catalogs, so that they could request additional information if needed.

Catalogs (or whatever material was available) were forwarded to Westat.

4.3.2 Sample Transcripts

Since transcript format varies greatly among school districts throughout the country, it was sometimes difficult to find the needed information on a transcript. This presented an obstacle to uniform treatment of information on transcripts. Another difficulty was encountered in determining the meaning of “coded” information found on some transcripts, particularly codes indicating the level of courses – that is, whether a course was honors or remedial level, or whether it was a special education course or part of another special program.

To solve this problem, Westat obtained sample transcripts of previous graduates, marked up to indicate where on the transcript the needed information was to be found, and how information regarding course level was coded. Westat requested three sample transcripts from each school: one containing honors level courses, one containing special education courses, and one “regular” transcript. Attached to each marked-up transcript was a Transcript Format Checklist, indicating the information to be marked and whether or not that piece of information was included on the school’s transcripts.

4.3.3 School Information Form (SIF)

The SIF was forwarded to Westat along with the other preliminary materials as described above. The SIF (see Appendix B) was completed by the field worker or a school staff member or sometimes by both. The name and position of the school’s HSTS coordinator who helped fill out the SIF appeared on the first page. The completed SIF contained information about the school in general, about sources of information within the school (if needed to complete HSTS data collection), about the course description materials, about graduation requirements and grading practices at the school, and about the format of the school’s transcripts. The field workers were instructed to fill out the SIF completely, or to indicate clearly on the SIF where the requested information could be found in the other materials provided by the school.

4.3.4 School Questionnaire

The School Questionnaire (formerly called the School Characteristics and Policies Questionnaire) (Appendix A) is a 54-item questionnaire that collected information about school, teacher,

and home factors that might relate to student achievement. It was completed by a school official (usually the principal) as part of NAEP for the NAEP participating schools.

4.4 Identifying the Sample Students and Obtaining Transcripts

The 1998 HSTS used the NAEP sample for selecting schools and students in NAEP-participating schools. For schools that participated in NAEP, the student sample was recorded on the NAEP Administration Schedules. For schools that did not participate in NAEP, the field worker drew a sample of students at the school. Our procedures for identifying students in schools with NAEP materials and in schools without NAEP materials are described in detail in separate sections below.

4.4.1 Schools with NAEP Materials

Schools that participated in NAEP identified students participating in the HSTS at the same time that the NAEP sample was selected. For all HSTS participants, a Disclosure Notice was placed in the student's cumulative record folder where it would be highly visible and thus make it easier to identify and collect needed transcripts after students had graduated.

Transcripts were requested for all students who were assessed, for sampled students who were absent during assessment, and for the SD/LEP students who were sampled but excluded by the school from participating in the 1998 NAEP assessment.

When graduation information was posted on transcripts (the date was provided by the school on the School Information Form), a field worker returned to the school to obtain the requested transcripts. For each NAEP school, the field worker was given a Transcript Request Form (TRF), Version 1 (Exhibit 4-7). In addition to Student Name and NAEP ID, it contained columns for entering graduation status, gender, birth month and year, race/ethnicity, SD status, LEP status, if receiving Title 1 services, and National School Lunch Program participation. Data available from NAEP files (NAEP ID and demographic variables) were preprinted on the form. The completed TRFs contained the following information:

- **Student Name** – The field worker recorded the first name, middle initial, and last name of each assessed, absent, or excluded student listed on the NAEP Administration Schedule. These entries were made to correspond to the preprinted NAEP ID.

Exhibit 4-7. Transcript request form – Version 1

1998 HIGH SCHOOL TRANSCRIPT STUDY TRANSCRIPT REQUEST FORM FOR SCHOOLS COOPERATING IN NAEF AND IN TRANSCRIPT

SCHOOL ID: 105-303-1
SCHOOL NAME: WEST HAVEN HIGH SCHOOL

STUDENT NAME	NAEP ID	EXIT	STATUS	SEX	BIRTHDATE	RACE/ETH	SD	LEP	TITLE I	NSLP	TRANSCRIPT RECEIVED
FIRST MI LAST											
	002-0910649			M	08/80	WHITE		N	N	N	4
	003-0900743			F	11/78	WHITE		N	N	N	4
	003-0901006			M	03/80	WHITE		N	N	N	4
	004-0903109			F	10/80	WHITE		N	N	N	4
	004-0904113			F	12/80	BLACK		N	N	N	4
	005-0904400			M	10/79	WHITE		N	N	N	4
	005-0904485			M	07/80	HISPANIC		N	N	N	4
	007-0905851			F	05/80	HISPANIC		N	N	N	4
	007-0909750			F	08/80	WHITE		N	N	N	4
	008-0900844			M	12/79	WHITE		N	N	N	4
	009-0907361			M	01/80	HISPANIC		N	N	N	4
	011-0901645			F	05/80	BLACK		N	N	N	4
	011-0904300			M	09/80	HISPANIC		N	N	N	4
	013-0960370			F	02/80	HISPANIC		N	N	N	4
	014-0913960			M	07/80	WHITE		N	N	N	4
	015-0901719			M	08/79	WHITE		N	N	N	4
	015-0914462			M	08/78	ASIAN/PI		N	N	N	4
	015-0914489			F	11/79	WHITE		N	N	N	4
	016-0900383			F	03/80	WHITE		N	N	N	4
	017-0914328			F	08/80	WHITE		N	N	N	4
	021-0933057			M	06/80	WHITE		N	N	N	4
	021-0933065			F	01/80	WHITE		N	N	N	4
	021-0933073			F	02/80	WHITE		N	N	N	4
	021-0933081			F	12/81	ASIAN/PI		N	N	N	4

- **NAEP ID** – The 10-digit NAEP assessment booklet numbers and SD/LEP questionnaire numbers for excluded students from the 1998 assessment were preprinted in ID order. This column on the TRF identified all students for whom transcripts were needed.
- **Exit Status** – Using information provided by the school, field workers assigned one of the following codes to describe each student’s outcome at the school:
 - a. Graduated with a standard diploma;
 - b. Graduated with an honors diploma;
 - c. Received a diploma with special education adjustments;
 - d. Received a certificate of attendance;
 - e. Still enrolled in this school;
 - f. Dropped out;
 - g. Other, such as transferred, GED, or unknown;
 - h. Out of scope; or
 - i. Completed course requirements but did not pass required tests.

Sometimes the exit status was determined directly from the transcripts and sometimes it was provided by other sources at the school.

- **Birthdate, Sex, and Race/Ethnicity** – Demographic information was generally preprinted for each sampled student. If not preprinted, it was recorded from the NAEP Administration Schedule. If the school informed a field worker that some of this information was incorrect, the field worker entered the correct information on the TRF.
- **SD and LEP Status** – For each student, it was recorded whether or not the student was classified by the school as SD and/or LEP.
- **National School Lunch Program (NSLP) and Title 1** – Yes or No for participation in each of these programs.
- **Transcript Received** – Field workers checked this column to document that the transcript for a given student had been received.

Once the Transcript Request Form was completed by carefully transferring student information from the Administration Schedules, the field worker filled out a summary box at the top of the form and requested transcripts according to the procedures set forth by the school. The Disclosure

Notice placed in students' folders at the time of the NAEP assessment helped to facilitate transcript collection in participating NAEP schools.

The following directions for completing the Transcript Request Form were given to the field worker.

1. Enter your name at the "Supervisor" line in the top box of the TRF.
2. Verify that the school has all of the pages of the Administration Schedules, comparing the school copies to your own (which were provided without names). Student names should be legible on the complete school copy.
3. Eliminate any non-12th graders by drawing a single line through their names.
4. Begin with the NAEP ID of the first student on the Administration Schedule. Find the corresponding NAEP ID on the Transcript Request Form. (These are printed in ID order.)
5. The birthdate, gender, race/ethnicity, SD/LEP, Title 1 status, and National School Lunch Program participation should all be preprinted on the TRF and should match the information recorded for that student on the Administration Schedule. If not, correct the information on the TRF after you have verified that you have matched entries correctly.
6. Record the student's full name from the Administration Schedule on the line of the Transcript Request Form with the same NAEP ID. Make a small check on the Administration Schedule as you go to indicate you have completed the transcription for a given student (this should be the last use of the Administration Schedule). In some schools, it may be necessary to record some form of school ID (e.g., Social Security Number) in addition to or in lieu of the student's name for the school to access the files. Make sure you're aware of this before you start completing the TRFs.
7. Continue this process for all 12th grade students on the Administration Schedules with one exception: any students who have been crossed off as "withdrawn" should be skipped in the process.
8. When you have gone through all of the Administration Schedules in this fashion, you should have a name entry corresponding to each NAEP ID preprinted on the TRF.
9. Code the "exit status" for each student at this time if it is available. Alternatively, this information may be recorded when the transcripts are collected. Confer with your school coordinator to determine the best way to get this information; it may not be on the transcript or it may be coded information.

Exhibit 4-8. Documentation of missing transcripts

NAEP School ID: _____

Supervisor: _____

Documentation of Missing Transcripts

Date: _____

School Name: _____

Number of Transcripts Requested: _____

Number of Transcripts Received: _____

of Regular Transcript: _____

of Honors Transcripts: _____

of Special Edu. Transcripts: _____

Reason(s) for Missing Transcripts: _____

[illegible]

10. Record the number of transcripts requested in the box at the top of the first page of the TRF. Record the number received at the time you obtain the transcripts. For each transcript received, place a checkmark in the “Transcript Received” column. Be sure to complete a “Documentation of Missing Transcripts” form (Exhibit 4-8) if you cannot obtain a transcript.

Once the field worker filled in the names of the students, some schools were able to access an electronic data file and copy the transcripts. In other schools, the transcripts were pulled from their folders and photocopied at the school.

Once the request was filled, the field worker reviewed the transcripts to ensure that a transcript was received for each 12th grade student who was selected for the NAEP assessment, whether or not that student had graduated. The field worker then checked each transcript for eligibility, understandability (e.g., are all the codes on it defined on the transcript or explained in the SIF?), and completeness and labeled each transcript with preprinted labels containing the School ID and the NAEP ID for the student. The field worker completed a Documentation of Missing Transcripts form to explain the reasons the school gave for any missing transcripts.

After the field worker collected and recorded all the information required on the sampled students and reviewed the transcripts for completeness and accuracy, he or she prepared the transcripts for transmittal to Westat. This procedure involved “masking” all personally identifiable information where it appeared on each transcript, using a broad felt tip marker or correction tape to line through or cover all identifiers.

Personal identifiers were also removed from the Transcript Request Forms. Before sending the TRFs to Westat, the field worker cut off the portion that included the students’ names, in order to comply with confidentiality provisions. The portion with the names was left in the school’s NAEP folder.

Schools were reimbursed at their standard rates for providing the transcripts. The field worker then completed a Shipping Transmittal Form (Exhibit 4-9) and returned it with the TRF, the transcripts, the Documentation of Missing Transcripts, and the SIF to Westat.

Exhibit 4-9. Shipping transmittal form

908842

1998 HSTS – SHIPPING TRANSMITTAL FORM
(INSTRUCTIONS: Fill out for each school and shipment)

School ID #: _____ School Name: _____

Supervisor: _____ School Shipment #: 1 2

Date Shipped: _____ Source of Sample: ☐ NAEP List
☐ New Sample

1. TRANSCRIPTS:

- 1) Total Number Requested _____
- 2) Number in This Shipment _____
- 3) Number Unavailable _____
- 4) Number to be sent/Estimated shipping date _____

IF SCHOOL DID NOT PARTICIPATE IN NAEP, COMPLETE THE FOLLOWING:

2. SD/LEP STUDENT QUESTIONNAIRES:

- 1) Total number requested _____
- 2) Number in this shipment _____
- 3) Number unavailable _____
- 4) Number to be sent _____

3. COURSE CATALOG: (check one)

- ☐ In this shipment
- ☐ To be shipped
- ☐ Unavailable

4. COURSE CATEGORY: (Check one for each year):

- | <u>1997-98</u> | <u>1996-97</u> | <u>1995-96</u> | <u>1994-95</u> |
|-------------------------------------------|-------------------------------------------|-------------------------------------------|-------------------------------------------|
| <input type="checkbox"/> In This Shipment | <input type="checkbox"/> In This Shipment | <input type="checkbox"/> In This Shipment | <input type="checkbox"/> In This Shipment |
| <input type="checkbox"/> To be Shipped | <input type="checkbox"/> To be Shipped | <input type="checkbox"/> To be Shipped | <input type="checkbox"/> To be Shipped |
| <input type="checkbox"/> Unavailable | <input type="checkbox"/> Unavailable | <input type="checkbox"/> Unavailable | <input type="checkbox"/> Unavailable |

5. SCHOOL INFORMATION FORM (Check one):

- ☐ In this shipment
- ☐ To be shipped

6. COURSE CATALOG CHECKLIST:

- ☐ In this shipment

7. TRANSCRIPT FORMAT CHECKLIST:

- ☐ In this shipment

4.4.2 Schools without NAEP Materials

In schools that did not participate in NAEP, the field worker first selected a sample of students, then requested transcripts for those students and followed the procedures described in the previous section for reviewing and shipping transcripts. The School Information Form was also completed, and course catalogs for the past four academic years were collected. The information in the catalogs was documented by completing the Course Catalog Checklist. At this point, the procedure was different. Rather than obtaining and annotating three example transcripts, as was done at the time of the NAEP visit to the school, the field worker used the Transcript Format Checklist to annotate three actual transcripts from among those that were collected.

In the schools that participated in HSTS but not in NAEP, the process of generating a sample of students began when the school produced a listing of all students who graduated from the 12th grade during the spring or summer of 1998. This list was requested during the preliminary call placed to the school when it was determined that the school would participate in HSTS. The following information was collected for each student selected for participation in HSTS:

- Exit status,
- Gender,
- Birthdate (month/year),
- Race/ethnicity,
- If student had a disability (SD),
- If student had limited English proficiency (LEP),
- If student was receiving Title 1 services, and
- If student was a participant in the National School Lunch Program.

These data were collected either with the list of 1998 graduates or after sampling, depending on which procedure was easier for the school. SD/LEP Questionnaires were not collected for students in schools that had not participated in NAEP.

Selecting the Sample

As described in Section 3.3, there were two basic sampling rules for the 1998 HSTS. These rules applied to all schools that required a new sample of students.

1. If there were 60 or fewer graduates listed, all graduates were included in the sample.
2. If there were more than 60 graduates listed, a sample of 50 students was drawn using a systematic random sampling.

Because the students in the HSTS-only schools did not have NAEP identification numbers, a set of IDs was preassigned for up to 60 students in each school. The Transcript Request Form—Version 2 (Exhibit 4-10) was preprinted with these IDs and had space for filling in each student's name and basic demographic characteristics.

The field worker, with the assistance of the school, completed the TRF and submitted it to the school staff. The transcripts were then provided to the field worker who reviewed and shipped them to Westat in the same manner as transcripts from schools participating in NAEP.

4.5 SD/LEP Questionnaire

Prior to 1996, the questionnaire that collected information from school staff about students with disabilities and students with limited English proficiency was called the IEP/LEP Questionnaire. It was retitled as the SD/LEP Questionnaire in 1996. The SD/LEP Questionnaire was completed for students sampled for NAEP and identified by the school as having a disability and/or for students with limited English proficiency. Westat asked the schools to have the person most knowledgeable about a student complete the questionnaire. In large schools, this person was typically a counselor, a special education teacher, or a teacher of English as a Second Language. In smaller schools, this person was typically a classroom teacher.

For schools participating in the 1998 NAEP, the SD/LEP Questionnaires were collected as part of the NAEP procedures. Questions one and two were used to determine which section(s) of the questionnaire should be completed. Part A (questions 3 through 19) was answered for a student with a

Exhibit 4-10. Transcript request form – Version 2

1998 HIGH SCHOOL TRANSCRIPT STUDY TRANSCRIPT REQUEST FORM FOR SCHOOLS COOPERATING IN NAEP AND IN TRANSCRIPT

SCHOOL ID: 103-302-1
SCHOOL NAME: DEDHAM HIGH SCHOOL

STUDENT NAME	NAEP ID	EXIT STATUS	SEX	BIRTHDATE	RACE/ETH	SD	LEP	TITLE I	NSLP	TRANSCRIPT RECEIVED
	990-0000001									
	990-0000002									
	990-0000003									
	990-0000004									
	990-0000005									
	990-0000006									
	990-0000007									
	990-0000008									
	990-0000009									
	990-0000010									
	990-0000011									
	990-0000012									
	990-0000013									
	990-0000014									
	990-0000015									
	990-0000016									
	990-0000017									
	990-0000018									
	990-0000019									
	990-0000020									
	990-0000021									
	990-0000022									
	990-0000023									
	990-0000024									

disability. Part B of the questionnaire (questions 20 through 41) was completed for an LEP student. If a student was classified as both SD and LEP, the entire questionnaire was completed. A copy of the questionnaire is included as Appendix C.

4.6 Sending Data to Westat

As with NAEP, safeguards were built into the procedures for the transcript study to ensure that applicable privacy requirements were met. These safeguards included the removal of all personal identifiers from the transcripts provided by the schools. When the transcripts left the school, students could be identified only by ID numbers. In schools where the NAEP information was available, the ID number was the same as the student's NAEP booklet number. In schools where a sample of students was drawn specifically for the HSTS, new IDs were generated.

After transcripts were collected and all information on sampled students recorded, field workers prepared the transcripts for transmittal to Westat. They first compared the data on the transcripts to the TRF to verify that they had obtained and correctly labeled the transcripts. At the same time, they noted on the TRF which transcripts were received and which were not. They then cut off the left hand column of the TRF, which contained the names of the students. The list of names remained in the schools (and was ultimately destroyed) and the remainder of the TRF was placed in the package to send to Westat.

The field workers masked all personally identifying information where it appeared on each transcript, using a broad felt tip marker to line through all identifiers. The types of personal identifiers and their location on the transcripts were different for each school and, sometimes, for the different categories of students within a single school. Field workers were careful to examine every transcript and line through the following information each time it appeared: student's name, parent's name, names of guardians or other relatives, addresses (including street, city, state, ZIP), phone numbers, and Social Security numbers or other student ID numbers.

A Shipping Transmittal Form accompanied all shipments to Westat and summarized the types and number of materials being sent. This form also gave information on whether the transcripts were from the NAEP list or a new sample and, if the school did not participate in NAEP, whether course catalogs and an SIF were included in the shipment.

4.7 Receipt and Review of Data from Data Collectors

When transcript study materials arrived at Westat, a receipt clerk carefully reviewed all items for accuracy and completeness. Transcripts were matched to the Transcript Request Form. Field workers were contacted immediately if further clarification was needed. Schools were reimbursed for the cost of producing the transcripts within two weeks of having their materials received at Westat.

An automated management system was developed and maintained at Westat. A disposition code structure was developed to indicate the status of each school's participation. As field workers reported the results of their contacts with district superintendents and individual schools, a receipt clerk keyed a disposition code for each school. Disposition reports were generated from the receipt system once a week so that home office staff could review the progress of securing cooperation from the sampled schools.

Once verified, information on the number of transcripts and course catalogs requested and received was entered in the receipt system by a data entry clerk. Weekly status reports were generated to monitor the progress of obtaining the transcripts. Transcripts and other school materials were maintained in individual school folders and stored until used by data preparation staff. Each school folder included the school's catalog or catalogs, Transcript Request Forms, student transcripts, Catalog and Transcript Format Checklists, a School Information Form, and a Shipping Transmittal Form.

Catalogs, sample transcripts, and SIF's were reviewed at Westat to ensure their completeness. Phone calls were made to the field workers or to schools, as needed, to resolve any questions regarding the content or accuracy of the materials.